Environmental Policy and the Choice of Abatement Technique: Evidence from Coal-Fired Power Plants[*]

Nathaniel O. Keohane
Yale School of Management

February 20, 2002

Abstract

This paper analyzes the impact of environmental policy instruments on the choice of pollution control techniques. I first recast the existing theoretical model of pollution abatement technology choice. Previous studies have found that the cost savings to a firm that adopts a technical innovation in pollution control are greater under a market-based instrument, such as an emissions tax or tradeable permit system, than under a command-and-control instrument such as an emissions rate standard. I show that this result is more general than previously recognized: in particular, it holds whether the new technique is more capital intensive (as in previous models) or less capital intensive. Hence, the choice of technique by firms should be more responsive to abatement costs under market-based instruments.

I then provide an empirical test of the theoretical predictions, using data from sulfur dioxide regulation of coal-fired electric power plants. My principal finding is that electric generating units subject to a tradeable permits regime were indeed more responsive to abatement costs than units under an emissions rate standard. That is, a given change in the average cost of scrubbing or coal-switching had a greater impact on the probability of scrubbing for units under the market-based regime. I also find that long-term coal contracts, regional location, and utility ownership class had significant effects on the choice of abatement technique.

1 Introduction

A central issue in the economic analysis of environmental policy is the optimal choice of policy instrument. In the short run, of course – that is, for fixed abatement technologies – market-based instruments such as emissions taxes and tradeable permits are cost-effective: in theory, they can achieve a given pollution control target at minimum cost. "Command and control" instruments, such as emissions standards, are in general not cost effective; a requirement that all firms meet a uniform standard will mean that marginal abatement costs vary among heterogeneous firms, so that overall costs could be reduced by reallocating pollution control responsibility.[1]

---

[*] This research was supported by generous fellowships from Resources for the Future and from the Graduate School of Arts and Sciences at Harvard University. Travel support was also provided by the Center for Energy and Environmental Policy Research at the Massachusetts Institute of Technology. Thanks are due to Denny Ellerman, Robert Stavins, Chris Avery, Jim Snyder, Ruben Lubowski, Richard Zeckhauser, and Adam Jaffe for helpful comments. Any remaining errors, of course, are mine alone.

[1] For rigorous development of these well-known results, see Baumol and Oates (1988).

Market-based policy instruments are also thought to provide advantages over command-and-control instruments over the long run. In particular, several theoretical studies have found that the incentive to adopt new, lower-cost abatement technologies is greater under market-based instruments, suggesting that those instruments will also be spurs to greater innovation and thus lower abatement costs in the long term. However, there has been little empirical analysis of these predictions. This paper studies pollution abatement technique choices by firms, in order to test the theoretical prediction that policy instruments differ in their effects on adoption of new technologies. Along the way, we recast the theoretical model of technology choice to make it somewhat more general and to provide a new, intuitive interpretation of the greater adoption incentive under market-based instruments.

An appropriate starting point for any discussion about technological change is the framework introduced by Schumpeter (1939). He distinguished three phases: invention (in which the creation of a new product or process is conceptualized); innovation (in which the new product or process is brought to market); and diffusion (in which the product or process is adopted by users in an industry or the economy.) We focus on the third step – more specifically, on the individual adoption decisions that together make up diffusion of a technology. The effects of environmental regulations such as emissions standards or taxes necessarily work by influencing the choices firms make about how to control pollution – that is, by influencing decisions about the adoption of abatement techniques. “Upstream” effects – for example, innovation of lower-cost abatement techniques – are indirect, mediated by the demand for new technologies that is created by the environmental regulation.

This paper first develops a simple theoretical model of abatement technique choice under different policy instruments, with an eye towards generating empirically testable hypotheses. The basic prediction from the theoretical model is that firms will be more sensitive to changes in abatement cost under a market-based policy (such as a tax or a system of tradeable allowances) than under an emissions rate standard. This conclusion agrees with well-known results in the existing literature. However, we show that the key result holds whether the new technology is more capital-intensive (as in previous models) or less capital-intensive than existing technologies. Our model provides a natural interpretation of the result in terms of the flexibility a polluting firm has, under a market-based instrument, to adjust its abatement under the new technology.

The heart of the paper is an empirical analysis of the use of flue-gas desulfurization devices, commonly known as “scrubbers,” to remove sulfur dioxide from the flue gases of coal-fired electric power plants. The history of sulfur dioxide regulation offers a rich context for empirical analysis of the effects of policy instruments on technique choice. Since 1970, federal regulation of sulfur dioxide (SO_2) has gone through three stages: an emissions rate standard, a technology standard, and a system of tradeable pollution allowances. In particular, the first and last of these policy regimes offer a testing ground for analyzing the impacts of policy instruments on technology adoption.

We show that the technique choices of units regulated by the tradeable permits system were indeed more sensitive to abatement costs than the choices made by units regulated by an emissions rate standard. That is, a given change in the average scrubbing cost or the average coal-switching cost had a bigger impact on the probability of scrubbing for units under the market-based regime. We also explore the effects of other possible influences on the decision of whether or not to scrub. In particular, we find that the presence of long-term coal contracts raised the likelihood of scrubbing, presumably due to increased costs associated with switching coals. We also find significant differences among regions and among ownership classes.

The absence of major market-based policy instruments in the real world has limited previous opportunities for empirical analysis.[2] In a study of home insulation, Stavins and Jaffe (1995) estimated the effects of energy taxes and adoption subsidies indirectly, using trends in energy prices and insulation cost. They concluded that energy taxes would have "noticeable impacts" on the insulating efficiency chosen by home builders; that adoption subsidies of the same magnitude would have even greater effects; and that building codes did not make a significant difference.

Bellas (1998) analyzed the costs of scrubbing at coal-fired power plants over time, looking for evidence of technological change.[3] Examining scrubbers built under the new-source performance standards of the 1970 and 1977 Clean Air Acts, Bellas failed to find any effects of scrubber vintage on cost.

Most of the work on the effects of environmental policy instrument choice on technology adoption has been theoretical rather than empirical. The predominant framework in the literature has involved what could be called the "discrete technology choice" model. A prime example of this approach is the paper by Milliman and Prince (MP) (1989). They consider the case of a firm, with some existing abatement technology, that contemplates adopting a certain new technology with lower marginal abatement costs. MP take the marginal abatement cost functions as given, and do not consider the output decisions of the firms.[4]

MP focus on the relative cost savings to an individual firm from adopting the lower-cost technology under different policy instruments. A key result is that the cost savings from the lower-cost technology are greater under a market-based instrument (such as a tax or system of tradeable permits) than under an emissions standard that is equivalent ex ante. MP attribute this result to the fact that under the market-based instrument, the firm increases its abatement when its marginal costs go down, while abatement is fixed under the standard. Hence under the market-based instrument the cost savings include savings not only on the abatement that was being performed under the old technology, but also savings on the increase in abatement (since by increasing abatement it pays less in taxes, or holds fewer permits). It follows that if there is a fixed cost to adopt the new technology, adoption should be more likely under the market-based instrument.

As a theoretical framework, this basic approach has several limitations. First, it restricts attention to the case in which a new technology has lower marginal costs, but comes at some initial fixed cost. This approach is appropriate for the case of a regulator requiring firms with existing abatement equipment to upgrade their controls.[5] However, in many real-world cases – such as the sulfur dioxide regulatory regimes discussed below – regulations are imposed on previously unregulated firms.

Other limitations are more fundamental. The MP approach starts by specifying marginal abatement cost functions, as if in a vacuum; the connections to the firm's output activity are ignored. The results depend on assuming fixed output and no entry or exit. Heterogeneity among

---

[2] For a comprehensive review of the literature on technological change and environmental policy, see Jaffe, Newell, and Stavins (forthcoming).

[3] In terms of Schumpeter's framework, Bellas sought to look at "innovation." We are focusing on studies of adoption and diffusion; we mention Bellas because of the relevance to the context of the present study. For a careful empirical analysis of "induced innovation," and a simulation of the effects of policy instruments on the direction of innovation in home appliances, see Newell et al. (1999).

[4] A similar modeling framework is used by Downing and White (1986); Malueg (1989); Jung, Krutilla, and Boyd (1996); and Fischer, Parry, and Pizer (1998). These models focus on other aspects of the effects of policy instruments and technological change, but they share the same underlying modeling framework described as the "MP approach."

[5] Note that this scenario also assumes implicitly that the existing technique either remains in place (to be upgraded) or, if replaced, has zero resale value.

firms is ignored or else not satisfactorily taken into account.[6] Moreover, the supply side is left unexplored; at most, the royalties collected by innovators are parameterized, but the structure of the supplying industry is not modeled.

At the same time, this theoretical literature has not derived concrete, testable hypotheses. Even starting from the existing framework, therefore, some work would be required to produce hypotheses appropriate for empirical analysis. For this reason, and because of the theoretical limitations just discussed, constructing a new framework from the ground up is a worthwhile endeavor. The theoretical model presented in this paper is a start. It does not address many of the problems above, but it does lay the foundation for a more general model. It also yields an explicit and testable hypothesis.

We develop a simple model of technique choice in the next section. Section 3 considers the model in the context of a particular area of environmental regulation – the control of sulfur dioxide from coal-fired electric power plants – and presents the econometric model to be estimated. Section 4 discusses the data used in the analysis, paying special attention to the construction of estimates for the cost of burning low-sulfur coal. Econometric results are presented and discussed in section 5. Section 6 concludes.

2 A theoretical model of technique choice

In this section we present a model of abatement technique choice, focusing on the following question: How does the choice of policy instrument affect the cost savings to firms from changes in pollution abatement techniques? We have in mind a setting where a polluting firm faces a choice between installing an old and a new version of a similar technology. To take an example from our later empirical analysis, we can imagine a firm choosing between one scrubber and another with slightly different costs.

Because our primary goal is to generate empirically testable hypotheses, the model is necessarily simple. It focuses only on the adoption of technical innovations in pollution control; it makes a number of simplifying assumptions about the output market; it ignores important aspects of adoption decisions at the industry level, such as the effects of strategic interactions among firms.[7] Moreover, it yields no direct conclusions about the welfare effects of different policy instruments.

Nonetheless, the model provides a useful framework for thinking about the effects of policy instruments on technique choice. The conclusions of the model are more general than those of the MP approach, and allow for a broader interpretation. The model also offers a foundation for future theoretical models of greater complexity and independent significance.

Consider a cost-minimizing firm producing output q and emitting E units of pollution. Let $m \equiv E/q$ denote the firm's emissions rate in units of pollution per unit of output, and let $\bar{m}$ represent the firm's emissions rate in the absence of any controls. To reduce emissions, the firm can reduce its output or reduce its emissions rate. These are complementary strategies; in general we would expect the firm do both. Here, however, we shall focus on the firm's choice of emissions rate. To simplify the analysis, therefore, we assume that output is fixed.[8]

---

[6] Although Jung, Krutilla, and Boyd (1996) attempt to take heterogeneity into account, their measure of a “market-level incentive” is of limited usefulness: incentives operate at the level of the firm, not the industry. Moreover, they ignore the effects of strategic interaction among firms. See the critique in Keohane (1999).

[7] These strategic effects are considered in Keohane (1999), in the context of a different model framework.

[8] An alternative assumption that would support our results is that reductions in the emissions rate are “separable”

Assumption 1 Firms produce a constant level of output $\bar{q}$; hence to reduce emissions they must reduce their emissions rate m.

We assume that reducing the emissions rate below $\bar{m}$ is costly; resources spent on emissions control are unavailable for other inputs to production. Viewing abatement efforts and emissions as two inputs into production, we can represent a given abatement technique by a "quasi-isoquant" giving the trade-off between the emissions rate m and resources spent on abatement c.[9] (See Figure 1.) Assuming that there are declining marginal returns to abatement efforts, this quasi-isoquant can be represented by a convex and decreasing function $c(m)$. On our assumption that output is fixed, the function $c(m)$ can also be interpreted as an emissions rate cost function in per-unit-output terms.

Letting k denote the capital cost per unit output,[10] we have $c(\bar{m}) = k$ and we can write the abatement cost function as $c(m) = k + \phi(m)$, where $\phi(m)$ represents the convex and decreasing variable abatement cost function.

Assumption 2 The cost of emitting at rate m, given initial (uncontrolled) rate $\bar{m}$, is given by $c(m; \bar{m})\bar{q}$; where $c(m; \bar{m}) = k + \phi(m; \bar{m})$, with $\phi' < 0$ and $\phi'' > 0$.

To simplify the comparison of capital and variable costs, consider a simple two-period model. In the first period, the firm decides which abatement technique to adopt, and pays any required capital cost.[11] In the second period, the firm emits at some chosen rate m, and pays the variable costs of abating from $\bar{m}$ to m. Without loss of generality, let the discount rate be zero, so that costs in both periods are weighted equally.[12]

We suppose that the firm faces a choice between two abatement techniques, denoted 1 and 2, with capital costs k_1 and k_2 and variable cost functions $\phi_1(m)$ and $\phi_2(m)$. As already noted, we can think of these techniques as "old" and "new" versions of some similar means of reducing pollution. If $k_1 < k_2$ and $\phi_1'(m) < \phi_2'(m); \forall m$; then technique 1 is unambiguously superior to technique 2. (Note that since $\phi_1(\bar{m}) = \phi_2(\bar{m}) = 0; \phi_1'(m) < \phi_2'(m); \forall m \Rightarrow \phi_1(m) < \phi_2(m); \forall m$.) In Figure 1, technique 1 is unambiguously superior to the technique represented by the dotted line. A more interesting case is the one in which technique 1, say, has a higher capital cost but lower marginal

---

from reductions in output, in the sense that the cost per unit output of achieving a given emissions rate is independent of output. That assumption would hold if, for example, the cost of reducing the emissions rate was solely a function of the percentage of emissions removed. However, the stronger Assumption 1 simplifies the model by allowing us to express capital costs as constants. As already noted, the assumption of fixed output has been made (either implicitly or explicitly) in the previous literature.

[9] Because we are expressing emissions and abatement cost in per-unit-output terms, the function $c(m)$ is not a true isoquant. On the other hand, we could draw a true isoquant relating total emissions E and the cost of abatement efforts.

[10] In this simple model, the capital cost is taken as given. In a more general model, the supply side of the market in pollution control technologies would be modeled; the capital cost would then correspond to the price of the technology, plus any fixed installation cost. In such a case, the capital cost k would clearly depend on the competitive structure of the supplying industry, among other things. As with many other important extensions of the model, we defer this to future work.

[11] An important limitation of this model is that it does not allow for the decision of when to adopt a new technique. Rather, the simple model we consider here — like the previous literature — contemplates a firm choosing among techniques at a given point in time. For a theoretical model of the decision when to adopt a given technology, see Jaffe and Stavins (1995).

[12] A change in the discount rate could equally well be represented as a change in the capital cost relative to the variable cost function.

costs for any emissions rate, so that $k_1 > k_2$ but $\phi_1'(m) < \phi_2'(m); \forall m$. This case is illustrated by curves 1 and 2 in Figure 1.

We compare two regulatory regimes: an emissions rate standard ("command and control") and a "market-based" policy which attaches a price to emissions. We couch the analysis in terms of a price on emissions equal to p. This price could be the equilibrium price of a tradeable emissions permit, or it could simply be a tax per unit of emissions. We assume that individual firms are "small," so that they take the permit price as independent of their choice of technique.[13]

Under an emissions rate standard s, the firm is required to emit at rate s, but pays nothing the pollution it emits; we assume perfect enforcement. Under the market-based instrument, the firm chooses its emissions rate but pays a price p on all units of emissions.[14] The firm chooses m to minimize its total costs, which are $c_j(m)q + pmq$ for technique j. Hence the firm chooses m_j^* such that $c_j'(m_j^*) = p$. Viewing the abatement cost function as analogous to an isoquant, p represents the price ratio between the emissions rate and abatement efforts (where the "price" of abatement effort is normalized to unity).

Since firms are assumed to be cost-minimizing, each firm will choose the technique that minimizes its total costs (abatement costs plus taxes or permit costs, if applicable). Let the relative cost savings under technique i (relative to technique j), measured per unit output, be denoted Δ_{j-i}: Under a standard, $\Delta^S_{j-i} \equiv c_j(s) - c_i(s)$; under a tax, $\Delta^T_{j-i} = c_j(m_j^*) + pm_j^* - c_i(m_i^*) - pm_i^*$.

To compare the two policy regimes, we need to make some assumption about how they are determined. A convenient assumption is that the policy regimes induce the same emissions rate ex ante.[15] We can imagine that before the start of the first period, at time 0, only one of the two technologies is available, and the regulator chooses the stringency of the market-based instrument or standard according to that initial technique.

Assumption 3 Whether the regulator chooses an emissions rate standard s or a market-based standard with emissions price p (inducing m_j^* such that $\phi_j'(m_j^*) = p$), the regulatory regimes are equivalent to one another under the technique j: $s = m_j^*$, where j is the initial technique.

Given this ex ante equivalence, a firm will find the newly introduced technique relatively more attractive under the market-based instrument than under the standard. This result holds true whether the newly introduced technique is more or less capital-intensive than the initial technique.

---

[13] Note that we are focusing on the comparison between a market-based instrument and an emissions standard. The key to our analysis is the presence of a price on emissions, rather than the magnitude of that price.

As shown elsewhere (Keohane, 1999), in equilibrium the fraction of firms in an industry adopting a lower-marginal-cost technique will be greater under a tax than under a tradeable allowances system that is equivalent under the ex ante technology. This is because the diffusion of the lower-cost technology throughout the regulated industry results in lower marginal costs for the industry as a whole, and thus a fall in the permit price. Because the adoption incentive is increasing in the price on emissions, this effect reduces the equilibrium incentive to adopt the new technology under the tradeable permit system, relative to the tax. The opposite result would hold true for a new technique with higher marginal costs and lower capital costs.

[14] If permits are auctioned, the firm will pay p on all its permits. If each firm receives some free allocation of permits from the government, p will still represent the cost of holding a permit on the margin, and thus the effect on the firm's abatement decision and cost savings will be the same.

[15] Note that this assumption strictly makes sense only if emissions standards are targeted to each firm, or alternatively if firms are homogeneous. However, it provides a natural "normalization" under which to compare the two policy instruments.

The general flavor of the result will carry over to heterogeneous firms with a uniform standard. For a discussion of notions of "equivalence" in that case, see Keohane (1999).

Proposition 1 Let Assumptions 1-3 hold. Consider a firm that faces a choice between abatement technologies 1 and 2, where $k_1 > k_2$ but $\phi_1'(m) < \phi_2'(m); \forall m$. If the initial technique is technique 1, then the cost savings to a firm from technique 2 relative to technique 1 will be greater under the market-based instrument. Likewise, if the initial technique is technique 2, the relative cost savings from technique 1 will be greater under the market-based instrument.

Proof: Consider the case where technique 1 is the initial technique. If the initial technique is technique 1, then under the market-based instrument the relative cost savings from using technique 2 is $\Delta^T_{1-2} \equiv c_1(m_1^*) + pm_1^* - c_2(m_2^*) - pm_2^* = c_1(s) + pm_1^* - c_2(m_2^*) - tm_2^*$ by A3. Substituting for $c_1(s)$ and $c_2(m_2^*)$, we have $\Delta^T_{1-2} = k_1 + \phi_1(s) - k_2 - \phi_2(s) + (\phi_2(s) + ps - \phi_2(m_2^*) - tm_2^*)$. Since m_2^* is the cost-minimizing level of m, $\phi_2(s) + ps - \phi_2(m_2^*) - pm_2^* > 0$. Hence $\Delta^T_{1-2} > k_1 - k_2 + \phi_1(s) - \phi_2(s) = \Delta^S_{1-2}$. By a similar argument, the relative cost savings from using technique 1 when the initial technique is technique 2 is given by $\Delta^T_{2-1} = \phi_2(s) + k_2 + ps - \phi_1(m_1^*) - pm_1^* - k_1 > k_2 - k_1 + \phi_2(s) - \phi_2(s) = \Delta^S_{2-1}$. QED

Figure 2a illustrates the proposition for the case in which the initial technique is technique 2, so that the new technique is more capital intensive. The two tangent lines are parallel, with slope equal to the emissions price p. By the assumption of equivalence, the standard s is the emissions rate at which the emissions price line is tangent to curve 2. With technology 1, the firm emits at rate m_1 under the market-based instrument.

The cost savings under the market-based instrument is given by the distance b between the two emissions price lines. Total costs would be found by adding the total tax bill or permits costs to the abatement cost, or equivalently by finding the vertical intercept of the tax line. (Note that the emissions price line plays the role of the isocost line in this framework.) Meanwhile, the savings under the standard is simply the distance between the cost functions at emissions rate s, which is the distance a < b. Figure 2b demonstrates the result for the case in which the new technique is the less capital-intensive technique 2. Again, the cost savings under the tax are given by b, which is strictly larger than a, the cost savings under the standard.

Because the previous literature has considered only the lower-marginal-cost case, interpretation of the results has focused on the greater abatement under the lower-marginal-cost technique as the source of the cost savings. Of course, that interpretation does not apply to the case in which a less capital-intensive technique is introduced.

A more general interpretation of the greater cost savings under the market-based policy centers on the flexibility allowed by such an instrument. Under a tax or a permit system, a firm can adjust its emissions to minimize its sum of abatement cost and tax. Our assumption of equivalence simply holds that the market-based instrument makes emissions rate s optimal under the initial technique. If we introduce a technical innovation and hold the emissions price constant, the innovation will always be relatively more attractive to the firm under the market-based instrument than under the standard, simply because under the market-based instrument the firm the can re-optimize along its new "quasi-isoquant."

Thus the greater cost savings under the market-based regime result from a kind of "substitution effect." Given an abatement cost function, the firm's cost-minimizing emissions rate equalizes its marginal cost of abatement and its marginal payment on emissions from doing so. A move from one cost function to another function with different marginal costs upsets this balance: on this new cost function, the firm finds it optimal to substitute abatement for emissions payments or vice-versa.[16]

---

[16] The result is also somewhat analogous to familiar results from consumer theory, such as the reason that Slutsky

Proposition 1 leads directly to a testable hypothesis: firms regulated by a market-based instrument will be more sensitive to the costs of abatement, in their choice of abatement technique, than firms regulated by a command-and-control instrument. By “more sensitive to the costs of abatement,” we mean that a given change in the underlying cost function – i.e., a given technological innovation – will have a greater impact on the firm's choice of technique. Because the cost savings from a technical innovation will be greater under the market-based instrument, a firm under that regime will be more likely to adopt the new technique, all else equal. Hence it will also be more likely to adopt that technique than some other alternative means of reducing pollution.

Figure 3 illustrates this result for the particularly simple case in which the technical innovation is a reduction in variable abatement cost at the same capital cost.[17] The standard s and market-based instrument (with emissions price p) are equivalent under the initial technology. For a given change in the adoption cost function, the cost savings under the market-based instrument (given by b) will be greater than under the standard (given by a).

In the empirical analysis below, we observe the average cost of abatement, and use that as our index of the cost function, inferring changes in the cost function from observed changes in average abatement cost. Note that this amplifies the effect of the policy instrument. The average costs of abatement are the slopes of the three rays from $(\bar{m}, 0)$ to the relevant points on the abatement cost functions. Under the standard, the firm's average abatement cost goes from AC_0 to $AC_1(s)$; under the market-based instrument, it goes from AC_0 to $AC_1(p)$. In the latter case, the firm responds to the change in the cost function by substituting abatement (which is now cheaper on the margin) for emissions payments. Thus the change in average abatement cost is smaller under the market-based instrument. However, the adoption decision is based on the comparison of the total cost of complying with the regulation. Under the emissions rate standard, compliance cost and abatement cost are identical; but under the market-based regime, the compliance cost is the cost of abatement plus the cost of emissions payments (the tax bill or the net cost of permits). As shown in Figure 3, the smaller change in average abatement cost under the market-based regime corresponds to a larger change in compliance cost.[18]

---

wealth compensation leaves consumers better off after a price change. Slutsky compensation allows the consumer to purchase her original bundle at the new prices. Because this allows the consumer to re-optimize to a higher indifference curve, it leaves the consumer better off than before the price change. In the case of the introduction of a new technique, the “price ratio” stays constant, while the isoquant shifts. Nonetheless, the underlying principle is essentially the same. Making the tax and standard equivalent under the initial technique plays the same role as allowing the consumer to buy the same bundle under the new price ratio.

[17] We use this case of an “unambiguous improvement” to keep the discussion and the figures simple. Because the result flows from Proposition 1, it would also apply for the general case of a change in technology. (The exception is for the opposite extreme case of a drop in capital cost without an increase in variable cost; it is easy to show that in this special case the cost savings will be identical under the two regimes, because the abatement level under the tax is unchanged by such a “lump-sum” change in abatement cost.)

[18] This is a general result. The fact that the ex post average cost under the market-based regime is higher is of course due to the fact that the average costs are increasing in the amount of abatement. But it is easy to show that given a price on emissions, a cost-minimizing firm will always abate (if it abates at all) in the region of increasing average abatement costs. This result is exactly analogous to the standard result that a profit-maximizing firm will only produce output in the region of increasing average costs.

3 The case of sulfur dioxide control

3.1 The context

Sulfur dioxide control by coal-fired electric power plants in the United States provides a ready testing ground for the model of adoption just outlined. Of course, we do not directly observe an individual unit responding to changes in abatement cost functions. However, we do observe a group of units with different abatement costs choosing different techniques. By estimating the effects of abatement cost on the choice of technique across units, holding other unit characteristics constant, we can indirectly estimate the effects of changes in technology on technique choice.[19]

Burning coal to produce electricity produces sulfur dioxide (SO_2) as a byproduct, because coal contains sulfur. In downwind urban areas, SO_2 contributes to respiratory ailments and morbidity. Injected into the atmosphere by tall stacks, SO_2 returns to earth as sulfuric acid in precipitation, and thus is a primary component of acid rain.

Government regulation of sulfur dioxide emissions at the federal level has undergone three phases. Under the Clean Air Act Amendments of 1970, new generating units were subject to a maximum allowable rate of SO2 emissions, but could choose how to meet that standard.[20] In the second phase, starting with units constructed after 1979, new sources were required not only to meet the prior emissions rate, but also to install flue-gas desulfurization devices, or "scrubbers."

Title IV of the 1990 Clean Air Act Amendments introduced a system of tradeable allowances for SO2 emissions, applicable to already existing units – those that had been grandfathered under the earlier policies. Phase I of the program started in 1995. It applied to the largest, dirtiest existing generating units, which had been exempt from the earlier new-source performance standards. Allowances are denominated in tons, and are distributed annually, roughly according to each unit's "baseline" emissions between 1985 and 1987. At the end of each year, each unit must submit a number of allowances equal to its emissions in that year.

Sulfur dioxide regulation thus presents us with three clearly defined policy regimes: a uniform emissions rate standard, a technology standard, and a system of tradeable permits. Because we are interested in the determinants of technique choice, the second regime is of little interest: units were required to install scrubbers. The first and third regimes, however, offer a test of the theory sketched above.

The manager of a coal-fired generating unit has two primary means of reducing the emissions rate of sulfur dioxide, or SO_2. First, he can switch to a coal with a lower sulfur content.[21] Switching coals does not require a large capital investment, but low-sulfur coal fetches a premium.[22]

---

[19] Note that we are not portraying scrubbing or coal-switching as a superior technique. Rather, we are using observed choices between scrubbing and switching as a means of measuring how sensitive the technique choice is to differences in abatement costs.

[20] In principle, this emissions rate standard applied to all units on which construction began between 1972 and 1979. Because of construction delays and political deals, the actual coverage extended to units built as late as 1991.

[21] Of course, for newly constructed units, under the NSPS-D regime, a unit is not "switching" coals, but rather burning a lower-sulfur coal than it would in the absence of regulation. For ease of discussion, we will use the term "coal-switching" to include the burning of low-sulfur coal by NSPS-D units as well as Title IV units.

[22] In fact, the existence of such a price premium will also depend on suitable conditions on the supply side and in other markets for the fuel. Loosely speaking, a price premium can arise out of a model of "vertical differentiation" in the market for the fuel, since all firms subject to the regulation will attach a higher value to the low-emissions fuel, all else equal. For present purposes, we simply assume the existence of the premium, leaving the exploration of its determinants for future research.

Alternatively, the manager can install a flue gas desulfurization device, or "scrubber" – a building-sized piece of equipment that removes sulfur dioxide from the flue gases, usually by reaction with a limestone slurry. It is an expensive capital investment. But because removal efficiencies are very high (up to 95%), the unit can continue to burn cheap high-sulfur coal.

3.2 The econometric model

Consider an individual firm subjected to one of the two policy regimes discussed above – a uniform emissions rate standard for new sources (NSPS-D) or a system of tradeable allowances (Title IV). We assume that the technique choice was made at a single point in time: when the unit is being built, in the case of NSPS-D plants, or in the year or two after the passage of the 1990 Clean Air Act, for Title IV plants.[23]

3.2.1 The anticipated costs of switching and scrubbing

We model the decision of whether or not to install a scrubber, because that decision is what we observe in the data.[24] We suppose that at the time of the scrubbing decision, the manager of each unit weighed the anticipated costs of one technique against the other in choosing a technique. We use the term "anticipated costs" to underscore the notion that the choice of technique is determined not by what the actual costs turn out to be, but rather by what costs are anticipated by the decision maker.

At the core of the manager's estimate of the anticipated costs are the capital and operating costs of the two techniques. Both the capital and operating costs of scrubbing may vary among plants, due to such variables as their size, the amount of coal they burn, the sulfur content of their coal, their location, and the age of their boiler (for retrofitted scrubbers). Denoting scrubbing by a superscript "s", let C_i^s denote the average costs of abating from $\hat{m}_i$ to m by scrubbing for unit i, where average cost includes both capital and operating cost, and is expressed per pound of sulfur dioxide abated.[25]

The costs of switching fuels are driven by the premium for low-sulfur coal. In part, this premium will be set by the demand for low-sulfur coal (a function of the type and stringency of regulation); this component would be constant across plants. A second component of the sulfur premium, however, is heterogeneous: it varies with location. Both power plants and coal are geographically dispersed. Because transportation costs make up a large fraction of the price of delivered coal, the effective premium for low-sulfur coal will be lower for those plants that are closer to low-sulfur coal deposits.

For similar reasons, the capital cost of switching will vary across plants. Coals from different regions differ along other dimensions than their sulfur content: heat content, ash content, and

---

[23] In theory, Title IV units could have chosen to install scrubber at any time. In fact, they either decided to install a scrubber for operation by early on in Phase I, or they did not install one at all. Thus, for our data, the assumption that the decision was made in 1991 or 1992 is appropriate.

Nonetheless, the fact that a Title IV unit had the option to delay investment in a scrubber needs to be taken into account. We return to this subject in our discussion of the econometric model, below. See also the discussion in note 10 above.

[24] It would also be of interest to model the coal-switching decision directly. Doing so, however, requires determining which units switched coals, and what coals they switched to. While we hope to address this issue in future work, for current purposes we focus on the presence or absence of a scrubber.

[25] We thus assume that the capital costs are appropriately annualized, so as to be comparable to operating costs.

"grindability" are particularly important. The greater the difference between the low-sulfur coal and the coal the boiler was designed for, the higher the capital costs to convert the boiler. Denoting coal-switching by the superscript "c", let C_i^c denote unit i's average combined capital and operating costs (per pound of sulfur dioxide abated) of abating from $\bar{m}_i$ to m by coal-switching, where the capital cost is appropriately annualized.

Other factors may also play a role in the relative costs of the two techniques, without affecting measured capital and operating costs directly. These include prior experience with scrubbing; statutory biases in state law that favor one technique or the other; regional location; the amount of coal under long-term contract; whether a plant is located at the mouth of a mine; and the ownership structure of the parent utility. To the extent that information about experience diffuses among plant personnel within a utility, we might expect past experience at the utility level to affect the probability of scrubbing at the unit level. Experience with installing earlier scrubbers might lower the total costs of installing a new one, making scrubbing more likely. Alternatively, if scrubbers at other units owned by the utility have proven more expensive than anticipated – for example, less reliable or harder to operate than expected – past experience might dissuade managers from installing scrubbers, making scrubbing less likely.

Statutory biases in state laws might also affect the true costs of scrubbing. In the wake of the 1990 Clean Air Act, a few states with high-sulfur coal reserves enacted legislation imposing requirements on in-state coal use or other requirements meant to favor the use of scrubbers.

Regional dummies may also pick up the effects of differences in state utility and environmental regulations. Northeastern states, for example, have more stringent air quality standards, which might well affect the probability of scrubbing. The "statutory bias" measure introduced above partially accounts for regulatory differences, but it focuses specifically on the attempts by a few states to discourage utilities from switching to out-of-state coal under Title IV.

The expected effects of long-term coal contracts are straightforward. Switching fuels would require breaking any existing long-term contracts; thus scrubbing ought to be more likely, the greater the amount of coal under long-term contract.

Minemouth plants are generally located off of railroad lines (and, of course, are not barge-served). Hence the costs of switching coals are likely to be significantly higher for minemouth plants than for rail- or barge-served plants. Moreover, we lack good estimates of sulfur price premia for minemouth plants, which are not observed buying coal on the spot market from other coal regions. Including a dummy for minemouth location corrects for the difference in sulfur premium and the difficulty in estimating it.

Finally, we allow for effects of the ownership class of the unit's parent utility (investor-owned, cooperative, municipal, state-owned, and federal) on the probability of scrubbing. We do this largely to account for the possibility that different ownership classes induce different mechanisms for technique choice. In particular, pressures to reduce emissions beyond federal requirements – whether for symbolic political reasons, or to help states meet ambient air quality standards – might make scrubbing more likely at publicly owned plants, and state-owned plants in particular.[26]

---

[26] Somewhat loosely, we may also conceive of such effects as changing the relative "total expected costs" of the two techniques; in this interpretation, the behavioral assumption remains that managers of all stripes minimize their costs, but the costs vary with ownership type.

3.2.2 The technique choice

If we had perfect information on the actual costs managers anticipated at the time of their decision whether or not to install a scrubber, analyzing technique choices would be a straightforward matter of comparing one total cost with the other. However, what we observe directly is limited, and in ways that affect our ability to estimate what is missing. Three problems arise. First, we do not observe the true "anticipated costs" a given manager faced. We can imagine each manager comparing two figures of anticipated average costs; we have information on the determinants of those cost estimates, and can generate our own estimates, but we do not know the actual figures themselves.

Second, the costs of fuel switching are not observed directly, because any given units uses only some of the coals available to it; units with scrubbers, in particular, are never observed buying low-sulfur coal. Thus we lack direct observations on price premia (and thus switching costs) for individual units. Instead, we estimate the price premia econometrically, as described in detail in section 4 below.

Third, and most importantly, the capital and operating costs of scrubbing are only observed for units that chose to install a scrubber. Moreover, our observations on scrubber cost are not random: we would anticipate that the units with scrubbers had lower scrubbing costs, relative to their switching costs. Hence there is a potential for selectivity bias. The problem is precisely analogous to the familiar problem in the labor literature studied by Heckman (1974). In analyzing the determinants of women's wages, he had to take into account the fact that women who chose to work were likely to have higher wages, or lower "reservation wages" (opportunity costs of not staying at home), than women who did not. Scrubber costs present the same type of problem. In using the observed costs to generate predicted costs for firms that did not choose to scrub, we must take into account the fact that our observations systematically come from units with relatively low costs of scrubbing.

To fix ideas, consider the following model of a unit manager's choice of technique, taking the information limitations into account. Suppose that the expected sum of average capital cost and average operating cost is a function of K observable variables. These variables include the unit's nameplate capacity, coal consumption, and coal sulfur content, as well as the age of the boiler in 1990 (for Title IV units) or the vintage of the scrubber (for NSPS-D units).[27] They also include dummy variables for regional location. Denote these variables by the vector $X = (X_1, X_2, \ldots, X_K)$, where $X_1, \ldots, X_{K_a}$ are continuous variables (such as generating capacity) and $X_{K_a+1}, \ldots, X_K$ are exponential terms with dummy variables, so that $X_{K_a+1} = e^{x_{K_a+1}}$, say, with x_{K_a+1} a dummy variable.

We then suppose that the average capital and operating costs of scrubbing for unit i are given by

$$C_i^s = \left(\prod_k X_{ki}^{\beta_k} \right) e^{\varepsilon_i} \tag{1}$$

---

[27] We do not include scrubber vintage for Title IV plants, simply because there was very little variation. For NSPS-D units, the scrubber vintage is the first year of scrubber operation for units with scrubbers, and of boiler operation for units without scrubbers. The law required that new sources meet the emissions rate standard continually; although some units evidently were able to delay operating the scrubber, a substantial majority of scrubbers went on-line when the associated boiler did. Thus the assumption that units which did not scrub would have operated a scrubber starting in the first year of boiler operation seems a reasonable one.

where the disturbance term ε_i is a mean-zero random variable representing the role of unobservable unit-specific elements of the scrubbing cost. We assume that $\varepsilon_i \sim N[0; \sigma_\varepsilon^2]$ and is independent across units.[28] In log-log form, writing $c_i^s = \ln C_i^s$ and $x_{ki} = \ln X_{ki}$, equation (1) becomes

$$c_i^s = \sum_k \beta_k x_{ki} + \varepsilon_i = x_i \beta + \varepsilon_i$$

Because we want to test for the effects of the policy regime on the scrubbing decision, and because we want to allow for changes in the cost of scrubbing over time, we allow each variable to enter in separately for each regime (i.e., as the product of interaction terms with dummy variables for NSPS-D units and Title IV units), and we include a dummy variable for Title IV units. Thus x_1 is the unit's nameplate capacity interacted with the Title IV dummy, with x_2 equal to nameplate capacity times the NSPS-D dummy. Similarly, the parameter vector β includes separate coefficients for key variables under different regimes.

Next, we assume that the effects of prior experience and statutory bias enter in multiplicatively. We also introduce a regime dummy that equals one for Title IV units, and allow for interactions between this dummy variable and the other two. We thus write unit i's anticipated average compliance cost of scrubbing as

$$(C_i^s)^{\theta^s} \cdot \exp\left(-\sum_j \delta_j d_{ji} + \eta_i^s \right); \qquad (2)$$

where again we have added a disturbance term η_i^s to allow for other unobserved influences on total scrubbing cost. We assume that η_i^s is normally distributed with mean zero and variance σ_s^2 and is independent across units. (We discuss possible correlation between η_i^s and ε_i^s below.) We return to the weighting parameter θ^s in more detail below.

Now consider the costs of switching coals. Let our estimate of the average capital and operating costs of coal-switching be denoted by $\hat{C}_i^c$, where the "hat" signifies that the coal-switching costs are econometric estimates, based on coal price data. In parallel to our assumptions on scrubbing costs, we assume that the influences of dummy variables for minemouth location and ownership class can be expressed in an exponential term, and we interact each with a regime dummy. We introduce a measure of coal under long-term contract in a similar fashion, including it in the exponential term; for convenience, let us include this continuous variable along with the dummy variables in the vector d. The anticipated average compliance cost of coal-switching for unit i is then given by

$$(\hat{C}_i^c)^{\theta^c} \cdot \exp\left(\sum_j \delta_j d_{ji} + \eta_i^c \right); \qquad (3)$$

---

[28] The assumption of normality is required for the application of the maximum likelihood estimation methods we use. It can be justified by viewing the error term ε as the sum of a large number of "small" idiosyncratic and unobservable disturbances. Because we are constrained to include in X only those variables which are observable for all units, including those that do not install scrubbers, a number of relevant variables – such as the design of the scrubber and sorbent used, the waste disposal method, and so on – are not in X but are likely to affect the costs. The assumption of normality amounts to an assumption that there are many such small disturbances, so that an application of the central limit theorem implies that the distribution of their sum is closely approximated by a normal distribution.

where the disturbance term ν_i^c represents unobserved influences on switching cost. We assume that ν_i^c is independent across units and is distributed as $N[0,\sigma_c^2]$.

The parameters $\theta = (\theta^s,\theta^c)$ offer measures of the sensitivity of the scrubbing decision to the estimated capital/operating costs of scrubbing and coal switching.[29] The flexibility they provide is important for three reasons. The first is practical. Recall that our measures of scrubbing and switching cost are econometrically estimated; thus including the parameters θ^s and θ^c offers flexibility in allowing for systematic errors in those estimates.

Second, the parameters θ allow for the possibility that managers weigh scrubbing costs and coal-switching costs differently. Such a difference might be the product of irrational behavior; but it might also be due to systematic differences in how state utility regulators treat the costs of the two abatement techniques, or how they treat the costs of capital investments versus operating cost (recall that scrubbing is a much more capital-intensive technique than coal-switching). If the costs of scrubbing and coal-switching are weighted equally by decision-makers, and if the estimated capital/operating costs of switching are equal (on average) to the estimates used by decision-makers, then $\theta^s = \theta^c$ for a given regime. While we might expect this equality to hold within a given regime, we include the parameters θ to allow the data to provide evidence of their magnitudes.

Third, and perhaps most importantly, the parameters θ allow for the possibility that in anticipating their compliance costs, managers weigh the abatement costs of techniques differently under different policy regimes. The theoretical model above predicts that the sensitivity to average costs of scrubbing and switching should be greater under the tradeable-permits system than under the emissions rate standard.[30] A comparison of these relative weights across regimes offers a test of that theoretical prediction.

A necessary condition for a unit to install a scrubber is that the anticipated average compliance cost of scrubbing is less than the anticipated average compliance cost of switching. Setting equation (2) less than equation (3), and taking logarithms of both sides, we can express this condition as

$$\theta^s c_i^s + \nu_i^s \le \theta^c \hat{c}_i^c + d_i\delta + \nu_i^c \tag{4}$$

$$x_i\beta + \varepsilon_i \le \theta^* \hat{c}_i^c + d_i\delta^* + \nu_i, \tag{5}$$

where $\hat{c}_i^c = \ln \hat{C}_i^c$, $\theta^* = \theta^c/\theta^s$; $d_i = (d_1,\ldots,d_J)$ is the vector of dummy variables for unit i, and $\delta^* = (1/\theta^s)\delta$ is the associated vector of parameters, divided by θ^s. The disturbance term $\nu_i \equiv (\nu_i^c - \nu_i^s)/\theta^s$ is normally distributed with mean zero; its variance is denoted by σ_ν^2.[31] Moreover, we allow it to be correlated with ε_i. The most likely source of such correlation is correlation between the capital/operating costs of scrubbing and the other components of scrubbing cost (i.e., between ν_i^s and ε_i), although we might also imagine correlation between the costs of scrubbing and the total costs of switching coals (i.e., between ε_i and ν_i^c). Thus ε_i and ν_i are distributed according to a bivariate normal distribution, given by $f(\varepsilon_i,\nu_i) \sim N[0,0,\sigma_\varepsilon^2,\sigma_\nu^2,\sigma_{\varepsilon\nu}]$:

---

[29] Note that without loss of generality we can assume that the weighting is applied to the whole “total cost” figure for switching, rather than just the capital/variable costs. In this case, we could interpret the parameter vector δ as measuring the “intrinsic” effect of dummy variables d, multiplied by θ^s or θ^c, with a similar adjustment to the error terms ν_i^c and ν_i^s.

[30] Recall that we are essentially using a unit's average cost of scrubbing as an “index” of, or proxy for, the cost function for that unit.

[31] The variance σ_ν^2, of course, incorporates the (unspecified) variances of ν_i^c and ν_i^s, along with any correlation between the two disturbances. Because we expect that θ^s varies between the policy regimes, the error term v_i will be heteroscedastic. Let $\sigma_*^2 = \mathrm{Var}(\nu_i^c - \nu_i^s)$. For NSPS-D plants, the disturbances will have variance $\sigma_\nu^2 = \sigma_*^2/(\theta^s_{NSPS})^2$. For Title IV plants, on the other hand, the variance will be $\sigma_*^2/(\theta^s_{TIV})^2$. We return to this issue in note -- below.

We refer to equation (4) as the "structural technique choice equation," since it is a necessary condition for the installation of a scrubber. For NSPS-D units, it is also a sufficient condition.

Under the tradeable allowances system of Title IV, however, we need to take two further factors into account. First, Title IV units had the option of choosing neither to scrub nor to switch, instead emitting at rate $\bar{m}_i$ and purchasing allowances to cover emissions. For those units, sufficient conditions for choosing to scrub are that the compliance costs of scrubbing (given by equation (2)) are less than those of switching, and that they are less than the anticipated allowance price. In effect, tradeable allowances impose a ceiling on the average cost of complying with the regulatory regime. To account for this "escape hatch" offered by allowances, we can distinguish between the effects of low switching costs (less than the allowance price) and high switching costs (greater than the allowance price). We expect that the choice of technique will be more sensitive to low switching costs than to high ones, since the relevant alternative to scrubbing in the latter case is buying permits, not switching coals.

Second, Title IV applied to existing firms, while the NSPS-D standards applied only to new sources. As noted above, we have modeled the technique choice as a one-time decision. For NSPS-D units, this is perfectly suitable. A unit subject to a new-source standard would choose one technique or the other at the outset, since it is typically more expensive to "retrofit" a scrubber onto an existing unit than to build it at the same time as the unit. On the other hand, Title IV units had the option to delay investment in a scrubber; they were retrofitting in any case. Because a scrubber is essentially an irreversible investment, and because there was considerable uncertainty surrounding the workings of the market for tradeable permits, such a "wait and see" strategy might have been attractive to many Title IV units.

If we had data on the levels of uncertainty that individual units had about future permit prices and coal prices, we could try to estimate the "option value" of waiting for each unit. However, we lack such data. Instead, we account imperfectly for the possibility of option value by including a dummy variable for Title IV units in equation (4).[32] If Title IV units perceived an option value from delaying investment in scrubbers, then the estimated coefficient on this dummy variable should be negative.

3.2.3 The econometric method

We cannot use equation (4) to estimate the determinants of the scrubbing decision without first generating estimates of scrubbing cost for units that chose not to scrub. Define the indicator variable $I = 1$ if a unit adopts a scrubber and 0 otherwise. Using our expression for scrubbing costs

---

[32] Although this is admittedly imperfect, it is not without reason. A simple model of a Title IV unit's decision would show that the option value would enter into the comparison of average compliance costs multiplicatively, as a fraction $\alpha < 1$, say, on the anticipated costs of coal-switching. Letting A_i denote firm i's compliance costs of scrubbing in equation (2), and B_i its costs of switching in equation (3), the condition for installing a scrubber taking option value into account would become $A_i \leq \alpha_i B_i$:

Because we lack information on the price expectations of individual units, we can only estimate a universal or "average" parameter α, rather than α_i. In the log-log form of our structural technique choice equation, we can represent this simply by a dummy variable on Title IV units, as we have done.

in equation (1), and denoting unobservable variables by tildes (~), we have the following model:

Cost equations

$ẽ^s_i$ = x$_i$¯ + "$_i$ (6)

$ẽ^c_i$ = µ¤ĉ^c_i + d$_i$±¤ + º$_i$ (7)

Selection mechanism

We observe c$_i$ and I$_i$, where

c$_i$ = $ẽ^s_i$ and I$_i$ = 1 if $ẽ^c_i$ ¸ $ẽ^s_i$

c$_i$ = 0 and I$_i$ = 0 otherwise.

Drawing on the treatment in Maddala (1983), we can analyze this model as follows. De¯ne z$_i$ = (x$_i$; d$_i$;c^c_i); the vector of all the explanatory variables in the model, and u$_i$ = ("$_i$ ¡ º$_i$)=¾, a random variable with a standard normal distribution, where ¾ = p¾$^2_"$ + ¾^2_o ¡ 2¾$_{"o}$:

The scrubbing condition $ẽ^c_i$ ¸ $ẽ^s_i$ can now be expressed as u$_i$ · z$_i$°, where ° is a vector of parameters normalized by ¾: The probability of scrubbing is now given by

Pr(I$_i$ = 1) = Pr(u$_i$ · z$_i$°) = ©(z$_i$°); (8)

where ©(¢) denotes the standard normal cumulative distribution function.

The econometric task is to generate estimates of scrubbing cost for all units, including those that did not scrub, and then use those estimates to analyze the structural equation (4). Because "$_i$ and u$_i$ are correlated, however, OLS regression of the cost equation (6) will produce biased estimates.[33] Essentially, we have a simultaneous equations model, where the two equations to be estimated are the cost equation (6) and the probability-of-scrubbing equation (8).

One estimation approach is to use a two-stage method: ¯rst estimating equation (8) by the probit method, then generating estimates of the hazard function Á(z$_i$ª)=©(z$_i$ª), and ¯nally running OLS with the estimated hazard function on the right-hand side to account for the selectivity bias. We use this method below to estimate the selectivity bias and demonstrate the applicability of the selection problem to this model.

This \two-step method" is ine±cient, however, so we use maximum-likelihood estimation to estimate equations (6) and (8) jointly. The likelihood function is

$$\text{L(¯; °;¾}^2_{"}\text{;¾}^2_{o}\text{;¾}_{"o}\text{) = }\underset{i}{\text{Y}}\ \text{·µZ}_{¡1}^{z_i°}\ \text{f}_{uj"}\text{(u}_i\text{jc}_i\text{ ¡ x}_i\text{¯)du}_i\text{ f}_{"}\text{(c}_i\text{ ¡ x}_i\text{¯)}\ \text{¶}^{¸I_i}\ \text{[1 ¡ ©(z}_i\text{°)]}^{1\text{ ¡ }I_i}\text{;} \qquad (9)$$

where f$_{uj"}$ is the conditional distribution of u$_i$ given "$_i$, and f$_"$ is the marginal distribution of ". Note that we can estimate ¯ directly, because it enters into the likelihood function independently

---

[33] The problem is that the expected value of "$_i$; conditional on observing y$_{si}$, is not zero. Rather "$_i$ given u$_i$ has a normal distribution with mean ¾$_{"u}$u$_i$ and variance ¾$^2_"$ ¡ ¾$^2_{"u}$, where ¾$_{"u}$ = (¾$^2_"$ ¡ ¾$_{"o}$)=¾ is the correlation between "$_i$ and u$_i$. Hence

$$\text{E("}_i\text{j u}_i\text{ · Z}_i\text{°) = ¡¾}_{"u}\ \text{µ}\frac{\text{Á(Z}_i\text{°)}}{\text{©(Z}_i\text{°)}}\text{¶;}$$

where Á(¢) and ©(¢) are the pdf and cdf of the normal distribution.

of σ, through the cost equation. Thus we can generate estimates of ln(average scrubbing cost), $\hat{c}_i^s = x_i\hat{\beta}$, for all units – including those that did not install scrubbers.[34]

With the cost estimates $\hat{c}_i^s$ in hand, we can then estimate the parameters of the structural equation (4) using the probit method. Define $\xi = (\varepsilon^s - \varepsilon^c)/\eta$, where $\eta^2 = \mathrm{Var}(\varepsilon^s - \varepsilon^c) = \sigma_s^2 + \sigma_c^2 - 2\mathrm{Cov}(\varepsilon^s, \varepsilon^c)$. We can write the probability of scrubbing as

$$\Pr(I_i = 1) = \Pr\left(\xi_i \leq \frac{\theta^c \hat{c}_i^c + d_i\delta - \theta^s \hat{c}_i^s}{\eta}\right) = \Phi\left(\frac{\theta^c \hat{c}_i^c + d_i\delta - \theta^s \hat{c}_i^s}{\eta}\right) = \Phi_i. \tag{10}$$

The corresponding likelihood function is

$$L = \prod_i \Phi_i^{I_i}(1 - \Phi_i)^{1 - I_i}. \tag{11}$$

Maximizing this function produces estimates of δ/η and θ/η, which is sufficient for the hypotheses we are interested in testing.

Finally, note that the structure of this model essentially imposes the use of a combined capital and operating cost measure. To estimate equations for capital and operating cost separately would require replacing equation (6) with two equations. Because the variables being estimated are logarithms, however, we could not simply add them to generate a selection equation comparable to the condition $\hat{c}_{ci} \geq \hat{c}_{si}$. That is, if we estimated the log of operating cost and the log of capital cost independently, we could not add those values together to get the log of the total scrubbing cost for comparison with the log of the switching cost. Since we measure costs in log-log form, therefore, separating the capital and operating cost in this selection model is not feasible.

This problem is not as significant as it may first appear. From a theoretical point of view, the analysis in section 2 demonstrates that because of the presence of fixed capital costs, average total costs of the various techniques are the critical determinants of technique choice. At a more practical level, the results presented in section 5 suggest that we are able to explain a great deal of the variation in scrubbing cost using the specification given here.

4 Data

4.1 Data on electricity generation and scrubber operation

The basic source of our data on electric power generation and scrubber operation is an annual survey conducted by the Energy Information Administration (EIA) of the Department of Energy. This survey, "Form EIA-767: Steam-Electric Plant Operation and Design Report" (EIA 1985-1998), gathers information on unit location, utility ownership, boiler and generator design and operation, and pollution abatement. We have data from 1985 to 1998; data for earlier years is not available.

---

[34] As noted above, the presence of θ^s in the denominator of σ_ν introduces the potential for heteroscedasticity in ν_i and thus in u_i. Heteroscedasticity is potentially a serious problem in maximum likelihood estimation, since so much rides on the proper specification of the distribution (Yatchew and Griliches 1984).

In the current context, however, those concerns are largely assuaged. Note first that the heteroscedasticity is an issue only in the estimation of the first-stage maximum likelihood equation. Moreover, our use of dummy variables in interaction terms in the cost-of-scrubbing equation means that the effect of each explanatory variable in x is estimated independently for each regime; thus the heteroscedasticity will not directly affect our estimates of the coefficients of the scrubbing cost equation. Indeed, the first-order conditions for maximizing the likelihood function are distinct for each regime.

Capital and operating costs for scrubbers are reported directly in the data. We have converted all costs to 1996 dollars. Capital costs were deflated using the Handy-Whitman public utility construction cost index (U. S. Census Bureau, 1972-1999). Scrubber operating costs (and all other costs, such as coal costs) were deflated using the "Intermediate materials, supplies, and components" producer price index figures compiled by the Bureau of Labor Statistics.[35] Capital costs were then annualized at a rate of 11.33%, following Ellerman et al. (2000). Average scrubbing costs (in cents per pound) were calculated by summing the annual capital charge and the average annual operating cost[36] and dividing by annual SO_2 abatement .[37]

Heat consumption is calculated by multiplying the average heat content of coal burned by the quantity of coal.[38] The sulfur content of coal burned is measured as the gross emissions rate the coal would produce, in pounds of sulfur dioxide per million Btus; it is calculated by taking into account the percent sulfur by weight of coal, the heat content of the coal, and the design characteristics of the boiler. For units with scrubbers, of course, we can use the actual sulfur content of the coal burned. But doing this for units without scrubbers would produce errant estimates of their likely scrubbing costs. For Title IV units, we use the sulfur content of the coal they burned just prior to the advent of regulation in 1990. For NSPS-D units, we use the sulfur content of the coal we estimate to be the cheapest coal for a given unit.

To take prior experience into account, we use a dummy variable that equals one if the parent utility of a unit had installed a scrubber on another unit at the time of the scrubbing decision. The dummy variable "Statutory bias" equals one for units in states identified by Lile and Burtraw (1998) as having a "statute bias toward capital-intensive compliance due to coal." In our sample, this corresponds to units located in Pennsylvania, Ohio, and Illinois. [39] Regional dummy variables follow the regional classifications used in the Handy-Whitman index.[40] The measure of long-term

---

[35] Downloaded from the Bureau's web site: http//:stats.bls.gov/datahome.htm.

[36] For NSPS-D units, we averaged the real operating costs over the first five years of operation, or 1985-89, whichever came last (recall that pre-1985 data is unavailable). For Title IV units, we averaged the real operating costs over 1995-98, excluding the first year of operation (since many Title IV units – unlike NSPS-D units – only operated their scrubbers for a fraction of the first year of operation.)

[37] Annual SO_2 emissions were estimated for the years corresponding to the annual operating cost (see previous note). We calculated SO_2 emissions on a mass-balance basis, using unit-level data on the sulfur and heat content and quantity of coal burned, the number of hours the scrubber was in operation as a fraction of the number of hours the boiler was in operation, and the unit-specific emissions factor (assigned using the design characteristics of the boiler and the characteristics of the coal burned). The process used was precisely the same process that would have been used, say, by regulators at EPA.

These estimated emissions were found to be consistent with data from EPA continuous emissions monitors, available for Title IV plants after 1995. Nonetheless, estimated emissions were used for both NSPS-D plants and Title IV plants for the sake of consistency.

[38] There is some potential for bias here, because we use the actual heat content of all units – both scrubbed and nonscrubbed. Some analysts have argued that the electricity generation (and thus the heat consumption) of scrubbed Title IV plants should be higher than that of nonscrubbed plants, all else equal. This is because we would expect the variable cost of scrubbing to be below that of switching, since switching involves such small capital costs. Thus the variable costs of electricity generation ought to be lower among scrubbed plants, and they should be dispatched at a higher rate. The data summarized in Table 2 provide some evidence of this claim, although one must also take into account the fact that nonscrubbed plants tend also to be systematically smaller (in capacity terms) than scrubbed plants.

[39] Utility ownership changes over time. We use the utility that owned the unit at the time of the adoption decision.

[40] Regional classifications by state are as follows (states which are not represented in our sample are in parentheses): North Atlantic: (ME),(VT),NH,(MA),(CT),(RI),NY,NJ,PA,WV,MD,DE; South Atlantic: (VA),NC,SC,KY,TN,GA,FL,AL,MS; North Central: OH,IN,IL,MI,WI,MN,MO,IA,KS,NE,(SD),ND; South Central:

coal contracts used was the tonnage of coal under 10-year contract at the plant level in 1990, as reported by the units on Form 767. Minemouth information was gathered from the coal data described below.

Tables 1 and 2 present descriptive statistics on the variables used in the econometric model. The dataset contains 248 Title IV units, of which 28 have scrubbers, and 167 NSPS-D units, of which 73 have scrubbers. We have observations on all variables, including scrubber costs, for 91 scrubber units (21 Title IV units and 70 NSPS-D units); thus these units are the basis for the first-step estimation of scrubbing cost.[41]

Note that the average scrubbing cost at Title IV units is much less than that at NSPS-D units.[42] At the same time, however, a smaller proportion of Title IV units chose to scrub: 21 of 220 (10%) versus 70 of 164 (43%). A lower proportion of scrubbers among units with lower average cost would appear to contradict the theoretical prediction that Title IV units ought to be more sensitive to cost. However, as the empirical results in the next section demonstrate, the theoretical prediction is borne out once we control for other differences among the units.

4.2 Estimating the cost of coal-switching

A key element of the decision whether or not to install a scrubber is the cost of switching to a lower-sulfur coal. The switching cost includes two basic components: the price premium a unit faces for lower-sulfur coal, and the capital investment required to convert a boiler to burn lower-sulfur coal. In this section, we discuss how the switching cost is estimated, considering each of these components in turn.

Table 3 lists the coal districts and the tonnage of coal sold on the spot market, the average SO_2 content[43], and the range of sulfur content from each district, for the period 1985-98. For reference, note that the emissions rate standard for the NSPS-D units was 1.2 lbs/mmBtu. Coal with lower sulfur content is known for obvious reasons as "compliance coal."

The variation in coal quality and sulfur content among coals from different regions is a key source of heterogeneity in sulfur dioxide abatement costs. Central Appalachia – southern and western West Virginia, eastern Kentucky, and western Virginia – is the primary coal-producing region in the U.S., providing nearly one-third of spot market coal over the fourteen-year period. Importantly, Central Appalachia is a source of both medium and low-sulfur coals. The major source of low-sulfur coal since railroad deregulation in the 1980s, however, has been Wyoming's Powder River Basin (PRB). Meanwhile, the Illinois Basin, comprising western Kentucky and southern Indiana and Illinois, is a prime source of high-sulfur coal.

LA,AR,TX,OK; Plateau: NM,AZ,CO,UT,NV,WY,MT,(ID).

[41] The three NSPS-D units dropped from the regression cost equation lack information on either capital or operating cost. Likewise, six of the seven Title IV units dropped also lack cost information.

One Title IV unit, Georgia Power's Yates 1 unit, is a dramatic outlier in cost and in sulfur content. It was an experimental unit set up to test a new scrubber design under a wide range of conditions; half of the funding came under a Department of Energy grant. In an interview, the manager of the scrubber project at Yates emphasized that the scrubber cost was roughly twice what it would otherwise have been, due to the need to gather data on scrubber performance for the DOE study. Indeed, including the Yates scrubber has a significant effect on the results. We treat the Yates unit as if it lacks cost information; effectively, it does. (No NSPS-D units are such significant outliers.)

[42] The difference is exaggerated somewhat in the means; the median scrubbing costs for units under Title IV and NSPS-D are 3.728 and 24.89, respectively.

[43] Here the SO_2 contents are calculated using standard "benchmark" emissions factors, along with the heat content and sulfur percentage of the coals.

Table 1: Summary statistics

Variable	Mean	Standard deviation	Minimum	Maximum
Title IV plants				
Scrubbed plants (N=21)				
Average scrubbing cost (cents/lb SO2)	4.835	2.835	2.183	11.51
Nameplate capacity (MW)	588.4	443.7	103.7	1300
Heat input (mmBtus)	3.78×10^7	2.70×10^7	8.58×10^6	8.96×10^7
Sulfur content (lbs SO2/mmBtus)	4.647	0.9734	2.618	5.863
Boiler age in 1990	20.90	7.543	11	36
Coal under 10-year contract (000 tons)	1848	2408	0	6600
Non-scrubbed plants (N=220)				
Nameplate capacity (MW)	326.0	223.6	75.00	1150
Heat input (mmBtus)	1.65×10^7	1.17×10^7	5.48×10^5	6.34×10^7
Sulfur content[a] (lbs SO2/mmBtus)	4.008	1.268	1.288	9.46
Boiler age in 1990	27.48	7.451	12	41
Coal under 10-year contract (000 tons)	603	1232	0	6600
NSPS-D plants				
Scrubbed plants (N=70)				
Average scrubbing cost (cents/lb SO2)	36.44	35.14	6.311	189.1
Nameplate capacity (MW)	485.8	176.3	173.0	913.8
Heat input (mmBtus)	2.77×10^7	1.26×10^7	7.88×10^6	5.65×10^7
Sulfur content (lbs SO2/mmBtus)	2.804	1.984	0.6821	6.751
Year scrubber began operation[b]	79.87	2.864	75	90
Non-scrubbed plants (N=94)				
Nameplate capacity (MW)	572.2	239.9	50.60	1300
Heat input (mmBtus)	2.73×10^7	1.27×10^7	1.85×10^6	7.06×10^7
Sulfur content[a] (lbs SO2/mmBtus)	2.517	1.653	0.6759	5.242
Year boiler began operation[c]	80.79	3.336	75	91

Notes:

[a] SO_2 content for nonscrubbed plants is SO_2 content of estimated cheapest coal.

[b] Last two digits of year scrubber began operation.

[c] Note that for nonscrubbed plants, this is last two digits of year boiler began operation.

Table 2: Frequencies of binary variables

Variable	Number of NSPS-D plants[a]	Number of Title IV plants	Total
Scrubbed plants[b] (82 NSPSD, 28 Title IV)			
Minemouth	24	0	24
Statutory bias (Title IV only)	0	6	6
Prior scrubber at utility	28	17	45
Regions:			
North atlantic	3	9	12
South atlantic	14	8	22
North central	25	11	36
South central 9	0	9	
Plateau	22	0	22
Ownership classes:			
Investor-owned	37	22	59
Cooperative	21	0	21
Municipal	4	4	8
State-owned	11	0	11
Federal	0	2	2
Nonscrubbed plants (101 NSPSD, 225 TitleIV)			
Minemouth	1	5	6
Statutory bias (Title IV only)	0	73	73
Prior scrubber at utility	87	90	167
Regions:			
North atlantic	5	41	46
South atlantic	16	64	80
North central	32	115	147
South central	35	0	35
Plateau	6	0	6
Ownership classes:			
Investor-owned	73	182	255
Cooperative	7	12	19
Municipal	11	2	13
Federal	0	24	24
State-owned	3	0	220

Notes:

[a] Entries are number of units for which corresponding dummy variable equals one.

[b] Includes scrubbed plants which lack cost information.

Table 3: U.S. coal regions

Spot market coal, 1985-1998

BOM District	Region	Number of Sales[a]	Million Tons[a]	Median Heat Content (Btu/lb)	Median SO_2 Content (lbs/mmBtu)	Range of Sulfur Content[b]
1	Northern Appalachia (PA, WV, MD)	17112	135.8	12370	2.7	1.6-3.5
2	Western Pennsylvania	7454	90.1	12510	2.7	1.9-4.4
3	Northern West Virginia	6094	72.6	12729	2.7	1.3-4.6
4	Ohio	8656	89.5	11803	5.2	2.8-6.7
6	Northern tip of WV	591	15.4	12194	5.9	4.3-6.7
7,8	Central Appalachia (WV, TN, KY, VA)	54543	574.9	12366	1.4	1.0-2.5
9	Kentucky	5376	109.8	11657	4.6	3.4-6.0
10	Illinois	5159	111.3	11523	4.1	1.9-5.7
11	Indiana	6380	108.7	11100	3.9	1.9-5.8
13	Alabama	2049	29.0	12103	2.3	1.1-3.7
17	Rocky Mtns. (CO, NM)	1452	34.8	11200	0.8	0.6-1.0
18	Southwest (AZ, NM)	219	15.1	9825	0.8	0.8-1.4
19[c]	So. Wyoming	579	21.9	10461	1.0	0.7-1.7
19[d]	Powder River Basin	6425	381.1	8577	0.7	0.5-1.0
22	Montana	775	33.5	9348	0.7	0.6-1.5

Notes: [a] Includes all observations in coal price dataset, not just observations for plants in coal price regressions.

[b] Range of middle eight percentiles (10th-90th pctile).

[c] District 19, excluding Powder River Basin.

[d] Includes counties Big Horn, Campbell, Converse, Lincoln, and Sheridan in Wyoming.

If data were available on the actual delivered prices of coal from each district over a range of sulfur levels, including the cheapest coal available to a given plant, calculating sulfur premia would be a straightforward matter of subtraction. In general, however, we do not observe plants buying low and high sulfur coal from Central Appalachia and low sulfur coal from the Powder River Basin, all at once. Moreover, NSPS-D units without scrubbers were constrained by regulation to burn low-sulfur coal, which may not have been the cheapest coal available to them. Hence we must estimate sulfur premia econometrically.

Two sources were used to generate the necessary data. Data on coal shipments was compiled from the Federal Energy Regulatory Commission's "FERC Form 423: Monthly Report of Cost and Quality of Fuels for Electric Power Plants" (FERC 1972-1998). Second, we gathered data for each power plant on the mode of transport (barge or rail, or neither for minemouth plants) and distance from major and local coal districts. This was done using the "1998 U.S. Coal Activity Map" (Resource Data International, 1998), the "Rand McNally Handy Railroad Atlas of the United States" (Rand McNally, 1982), and the "Rand McNally United States Road Atlas" (Rand McNally, 1994).

Using these data, price regression equations were estimated for coal from each district, as a function of distance, form of transportation, and coal quality.[44] The Appendix describes the process in more detail, and presents sample coal price regression equations.

These regression equations were used to generate predicted coal prices at the plant level[45] for coals from appropriate districts using the average sulfur content of each district (for plants without scrubbers) or the actual sulfur content observed (for plants with scrubbers).[46] A "cheapest coal" was identified for each plant.

Next, we estimated prices for low-sulfur coal, again for each plant.[47] Four areas are major sources of low-sulfur coal: Central Appalachia, the Powder River Basin, BOM District 1 (northern Appalachia: southwestern Pennsylvania, northeastern West Virginia, and western Maryland), and the Central Rockies area (BOM District 17) in western Colorado. The price premia were calculated simply by subtracting the price of the cheapest coal available from the price of the compliance coal. In some cases – plants in western states close to low-sulfur coal deposits – the compliance coal was the cheapest coal available, and hence the price premium was zero.

As noted above, a second component of the costs of switching fuels (for Title IV units) is the capital cost of converting a boiler to burn lower-sulfur coal. This capital cost depends on the

---

[44] A word on expectations is in order here. Our aim is to include in the data the coal prices that would have been relevant to the unit managers at the time they decided whether to scrub or switch coals. This inevitably raises the issue of what their expectations were. Did they rely on past price trends? Or did they have perfect foresight about future coal prices and operating costs? We choose a middle course. Some foresight seems reasonable: a manager choosing a technique in the early 1980s would have looked forward to the likely conditions at the end of the decade. On the other hand, it seems heroic to suppose that managers of NSPS-D units in the early '80s would have predicted the fall in transportation costs, or the rise in demand for low-sulfur coal induced by Title IV. For Title IV units, we used the estimated coal prices for 1995-1998. For NSPS-D units, we elected to use the same time periods as were used for calculating average scrubber costs: namely, 1985-1989 for units that began operating in 1985 or before, and the first five years of operation for all other units.

[45] The coal data is given by plant, not by unit. This is perfectly consistent with using it to derive coal prices, since one would expect prices to be identical for every unit at a plant. We rely on Forms EIA-767 for data on the characteristics of coal burned by each unit.

[46] This sulfur content corresponds to $\hat{m}$ in the theoretical model. Note that we are assuming that units with scrubbers chose the cheapest coal available.

[47] For our purposes, we define "low-sulfur coal" as compliance coal.

differences in quality between the high- and low-sulfur coals used. A unit that switches from high-sulfur Central Appalachian coal to low-sulfur Central Appalachian coal faces very small capital costs. On the other hand, if the same unit switched to Powder River Basin coal – a coal with very different characteristics from eastern coal, quite apart from sulfur content – a much larger capital investment would be required to adapt the boiler to the new coal.

Drawing on surveys performed by Denny Ellerman and his colleagues,[48] we use an estimate of fifty dollars per kilowatt of capacity ($50/kW^e) for conversion to western coal, and ten dollars per kilowatt capacity ($10/kW^e) for conversion to eastern coals. These conversion costs were converted into cents per mmBtus for each unit, using the unit's heat consumption.

We then estimated the coal-switching costs, in cents/mmBtus, by summing the estimates for the price premium and the conversion cost at each unit. For each unit, we estimated the cost of switching from the cheapest coal available to each relevant low-sulfur coal.[49] We then used the lowest such cost for each unit as our estimate of that unit's coal-switching cost.

Table 4 summarizes our estimates of switching costs, by regime and region. Note that coal-switching costs are generally lower among NSPS-D units. This is due to the fact that those units do not have to retrofit their boilers; hence their "switching costs" do not include the capital cost of converting a boiler. Note also that the switching costs of western (Plateau) units are much lower than of units in other regions: low-sulfur coal is relatively much closer than high-sulfur coal for those units, so that the effective sulfur premium is much lower.

Finally, we distinguish between "high" and "low" switching costs. If the estimated cost of switching was less than scrubbing, we label the average switching costs "low" switching costs.[50]

Note that the estimates of conversion costs are in cents/mmBtus, unadjusted for the drop in emissions achieved by switching coals. Because we are comparing switching costs with average scrubbing costs, we need to divide by the difference in sulfur content between the high- and low-sulfur coals for a given unit to get figures in cents/lb SO_2, comparable to the scrubbing costs. Equivalently, because we take logs of the average switching cost in our model equations, we can simply introduce the natural log of the difference in sulfur content as an explanatory variable in our model. We take the second approach in the analysis below. The new variable is ln(difference in SO_2 content).

5 Empirical analysis

5.1 Two-step method

We first consider the results of the two-step method suggested in various forms by Heckman (1976) and Lee (1976). This approach provides consistent estimates of the parameters, although it is not efficient. The method (see Maddala (1983) for an exposition) is to estimate equation (8) by the probit method. The resulting parameter estimates, the ^, are then used to generate estimates

---

[48] Ellerman, personal communication.

[49] For example, for rail-served units in the northwestern corner of Indiana, we estimated switching costs for Central Appalachian coal, Powder River Basin coal, and Central Rockies coal.

[50] We used the average allowance price for the years 1995-1998 to calculate the corresponding cost (in cents/mmBtus) of buying allowances to cover a pound of sulfur dioxide emissions. Given the average allowance price of $118.70 per ton, the cost of emitting a pound of sulfur dioxide was 5.93 cents.

The comparison between the costs of switching and allowance costs took into account the amount of abatement that could be achieved by switching coals.

Table 4: Coal-switching costs

Estimated costs, in cents/mmbtus

Unit regime and region	Mean	Standard deviation	Minimum	Maximum
Title IV plants				
Scrubbed plants				
North atlantic (9)[a]	24.8	4.73	19.2	32.1
South atlantic (8)	27.7	13.2	9.77	45.9
North central (11)	21.6	9.28	10.7	34.8
Total (28)	24.4	9.48	9.77	45.9
Nonscrubbed plants				
North atlantic (41)	28.6	12.8	0[b]	53.3
South atlantic (64)	24.7	13.9	7.81	58.0
North central (115)	23.5	14.6	7.50	62.8
Total (220)	24.8	14.1	0	62.8
NSPS-D plants				
Scrubbed plants				
North atlantic (3)	26.9	4.81	24.1	32.4
South atlantic (16)	19.7	15.4	0	41.1
North central (25)	17.3	19.2	0	58.8
South central (9)	27.6	41.4	0	82.9
Plateau (22)	0.441	1.39	0	4.64
Total (75)	14.7	21.3	0	82.9
Non-scrubbed plants				
North atlantic (5)	27.1	6.98	17.8	34.5
South atlantic (16)	15.2	13.7	3.66	37.9
North central (32)	12.8	18.8	0	70.1
South central (35)	9.85	34.3	0	98.0
Plateau (6)	1.42	2.21	0	4.64
Total (94)	12.1	19.8	0	98.0

Notes:

[a] Number of units in parentheses.

[b] Estimated switching cost is zero when compliance coal is cheapest coal.

Table 5: Second-step OLS regression results
Dependent variable is ln(average cost of scrubbing)

Parameter	Variable	Coefficient		Standard error[a]
β_1	Title IV ln(nameplate capacity)	0.207		0.339
β_2	NSPS-D ln(nameplate capacity)	0.062		0.290
β_3	Title IV ln(heat input)	-0.533		0.365
β_4	NSPS-D ln(heat input)	0.212		0.259
β_5	Title IV ln(SO_2 content)	-0.659		0.450
β_6	NSPS-D ln(SO_2 content)	-0.404	**	0.092
β_7	Title IV Boiler age in 1990	-0.012		0.016
β_8	NSPS-D Year operation began	-0.0002		0.0001
β_9	Title IV North Atlantic	0.045		0.254
β_{10}	Title IV South Atlantic	-0.419		0.249
β_{11}	NSPS-D North Atlantic	-0.286		0.238
β_{12}	NSPS-D South Atlantic	-0.042		0.135
β_{13}	NSPS-D South Central	-0.425	*	0.169
β_{14}	NSPS-D Plateau	0.716	**	0.143
β_{15}	Title IV dummy	8.93	*	3.84
$\sigma_{\varepsilon u}$	Selectivity bias term	-0.176	**	0.088
β_0	constant	0.206		2.03

Notes: * significant at 5 % level ** at 1 % level
[a] Standard errors corrected for estimation of the hazard function.

N = 91
$R^2 = 0.88$
F[16,74]=35.20

of the hazard function $h_i = \phi(z_i\Psi)/\Phi(z_i\Psi)$ for each unit. The negative of this hazard function is included in a OLS regression of scrubbing costs on the explanatory variables x, correcting for the selectivity bias. While we perform full maximum-likelihood estimation below, the preliminary two-step estimation serves two purposes: it provides a direct estimate of the effect of "selectivity bias," and it gives us some sense of the ability of our explanatory variables to explain scrubbing costs in a conventional regression framework.[51]

Table 5 presents the results of the second-step OLS regression of scrubber costs on the variables in x, which is of most relevance here. The standard errors have been corrected to allow for the fact that the hazard function is estimated, following Maddala (1983).[52] Two points are of interest from

---

[51] While an estimate of the selectivity bias can also be derived from the maximum likelihood estimation below, the two-stage method provides a more direct and convenient measure: the estimate of $\sigma_{\varepsilon u}$ comes straight out of the regression equation.

[52] The corrected variance matrix is given by $V = \sigma^2_\varepsilon (G_1'G_1)^{-1} - \sigma^2_{\varepsilon u}(G_1'G_1)^{-1}G_1'B(G_1'G_1)^{-1}$, where $G_1 = (X_1; -H_1)$, where X_1 is the matrix of all the observations (for scrubbed units) on the explanatory variables in the cost regression and H_1 is the vector of hazard terms $\{h_i\}$ for those units. The second term, the correction to the standard errors, involves the matrix $B = D_1 - D_1Z_1(Z'\Lambda Z)^{-1}Z_1D_1$, where Z is the matrix of all observations on

Table 5. First, the coefficient on the selectivity bias term is negative and significantly different from zero at the 5% significance level. That result confirms the presence of the "selectivity bias" discussed in section 3: the expected costs of scrubbing conditional on observing a scrubber are different than the unconditional expected costs. The sign of the coefficient implies that the units which adopted scrubbers had systematically lower scrubbing costs than those that did not.[53]

Second, the regression demonstrates that despite the use of "generalizable" explanatory variables in the cost regression (i.e., variables for which we have observations for all units, not just scrubber units) we are able to explain a significant amount of the variation in scrubbing costs. The R^2 of the regression is 0.88, and an F test that the coefficients equal zero is rejected at any level of significance.

The results presented in Table 5 may appear to suggest that the explanatory power of the regression is concentrated in a few variables – in particular, the regional dummies and the time trend for NSPS-D plants. But this ignores the role of the hazard term in the regression. The significance of the hazard term implies that the cost-of-scrubbing equation and probability-of-scrubbing equation are properly viewed as a system of simultaneous equations. Unit characteristics turn out to have strong effects in the first-step (reduced form) technique choice equation; thus they play a major role in determining the hazard function, which in turn drives the cost-of-scrubbing equation. We return to this issue in the context of maximum likelihood estimation below, where we test the hypothesis that the coefficients on the unit characteristics are zero.[54]

5.2 Maximum likelihood estimation

Although the two-step method just described provides consistent estimates of the parameters, maximum likelihood is preferable. As discussed above, we first perform ML estimation on the "reduced form" equation (9). Then, using the parameter estimates for the cost-of-scrubbing equation, we estimate scrubbing costs. We then use those costs in estimating the parameters of the "structural" technique choice equation given by (4). In describing our results, we refer to the estimation of the likelihood function (9) as the "first stage" estimation, and the subsequent estimation of the structural equation (11) as the "second stage." Note that these steps do not correspond to the two steps of the Heckman method described above: both steps of the latter method are contained in the first-stage maximum likelihood estimation below.

5.2.1 The reduced-form technique choice equation and cost-of-scrubbing equation

Using the likelihood function given by (9), the log-likelihood function to be maximized is

---

all the explanatory variables in z and Z_1 is the corresponding matrix for the units with scrubbers, D_1 is a diagonal matrix with terms $h_i(h_i + z_i\gamma)$ for scrubbed units, and Λ is a diagonal matrix of $h_i/(1 - \Phi_i)$ terms for all observations.

[53] Recall that the bias term equals $-\phi_i/\Phi_i$; this term is larger (less negative), the greater is the estimated likelihood of adoption. Strictly speaking, of course, $-\phi_i/\Phi_i$ is not monotonic in its argument; but for probabilities of adoption greater than one-half, it is. The sample means of the hazard term in our data are -3.63 for nonscrubbing plants and -0.365 for scrubbing plants.

[54] In fact, the concern that unit characteristics appear to be playing a minor role in determining scrubbing cost would be largely moot, even if true. We are less concerned here with the determinants of scrubber cost than with the effect of estimated costs on the adoption decision. It should not be surprising that in collapsing our cost information to one observation on average costs per unit, and then restricting ourselves to variables that are observed for non-scrubbed units, our parameter estimates suffer from imprecision.

$$\ln L = \sum_1 \left\{ \ln \int_{-\infty}^{z_i\gamma} f_{u|\varepsilon}(u_i|c_{si} - x_i\beta)\,du_i + \ln[f_\varepsilon(c_i - x_i\beta)] \right\} + \sum_0 \ln[1 - \Phi(z_i\gamma)] \qquad (12)$$

where $\sum_1$ is the summation over observations i with $I_i = 1$. The maximum likelihood results are provided in Tables 6a and 6b.[55]

Table 6a presents the parameter estimates for the first-stage technique choice equation. The γ can be thought of as "reduced form" parameters. Because we are more interested in the "structural" parameters in equation (4), we do not focus on the interpretations of the γ.

Table 6b presents the estimated parameters of the scrubbing cost equation, $\hat{\beta}$.[56] Although most of the coefficients are not precisely estimated, their signs are reasonable. For example, the estimates of β_1 and β_2 are both positive; this is what we would expect, since scrubbers on larger units tend to be more expensive (and we have controlled for sulfur content and heat consumption). Moreover, $\hat{\beta}_3$; $\hat{\beta}_5$; and $\hat{\beta}_6$ are all negative (although $\hat{\beta}_4 > 0$), which is also reasonable. Greater heat consumption for given sulfur content, or vice versa, implies higher gross SO_2 emissions. Because the costs of scrubbing increase with the percentage of emissions removed, greater gross emissions should imply a lower cost for a given tonnage of abatement.

Boiler age appears to have very little effect on cost. The time trend for NSPS-D plants is negative, but not significant at the 5% level.[57]

The estimates for the coefficients on the regional dummies suggest that costs are significantly lower in the South Central region, and significantly higher in the Plateau states. (The omitted dummy is for the North Central region.)

The parameter $\hat{\beta}_{13}$, the coefficient on the Title IV dummy in the scrubber cost equation, is large and positive, although not significantly different from zero. This suggests that any improvements in scrubbers that might have occurred between NSPS-D scrubbers and Title IV scrubbers is outweighed by the effects of greater abatement by Title IV units.

As noted above, the imprecision of the parameter estimates $\hat{\beta}$ in the cost equation in Table 6b raises the issue of the role of unit characteristics such as generator capacity and coal sulfur content in predicting scrubbing cost. A direct test is possible in the maximum-likelihood context. Because the scrubbing and scrubber cost equations comprise a simultaneous system, the relevant

---

[55] These were produced using the heckman command in Stata, which estimates the log-likelihood function in the form $\ln L = \sum_1 \left\{ \ln \Phi\left(\frac{z_i\gamma + (c_i - x_i\beta)\sigma_{\varepsilon u}/\sigma_\varepsilon^2}{\sqrt{1 - \sigma_{\varepsilon u}^2/\sigma_\varepsilon^2}} \right) - \frac{1}{2}\ln(2\pi\sigma_\varepsilon^2) - \frac{1}{2}\left(\frac{c_i - x_i\beta}{\sigma_\varepsilon}\right)^2 \right\} + \sum_0 \ln(1 - \Phi(z_i\gamma))$. This is equivalent to the log-likelihood function in the text, since $u_i|\varepsilon_i \sim N\left(\frac{\sigma_{\varepsilon u}}{\sigma_\varepsilon^2}\varepsilon_i;\ 1 - \frac{\sigma_{\varepsilon u}^2}{\sigma_\varepsilon^2}\right)$.

[56] The coefficients and standard errors do not reflect two potential sources of bias: the heteroskedasticity due to the variation in μ^s among regimes, noted above, and the fact that the estimates of switching cost are econometrically estimated rather than directly measured. The former problem is more significant than the latter. We return to it below.

[57] Nonetheless, the negative sign on the coefficient of vintage is worth noting, because there is a bias in the data which should tend to make the coefficient positive.

As already noted, our data begins in 1985. For units that began operation after 1985, the average cost figure is for the first five years of operation. For units that operated prior to 1985, however, the average cost figure used is from 1985-1989. This discrepancy matters because there is evidence that average costs of scrubbing fall over time, at least within the first several years after operation commences. This trend was noted by Bellas (1998) and is present in the data used here.

If average costs do fall over time, then our average cost measure would systematically underestimate the costs of early-vintage NSPS-D scrubbers relative to later ones. This error would tend to to make β_8 positive.

Table 6a: First-stage maximum likelihood estimates
Reduced-form technique choice equation
Dependent variable is $I_i = 1$ for scrubbed units, $I_i = 0$ otherwise

Parameter	Variable	Coefficient		Standard error
γ_1	Title IV ln(nameplate capacity)	-10.3	**	2.44
γ_2	NSPS-D ln(nameplate capacity)	-1.21		0.856
γ_3	Title IV ln(heat input)	9.86	**	2.30
γ_4	NSPS-D ln(heat input)	1.37		0.781
γ_5	Title IV ln(SO_2 content)	1.11		1.06
γ_6	NSPS-D ln(SO_2 content)	0.404		0.399
γ_7	Title IV Boiler age in 1990	-0.131	*	0.067
γ_8	NSPS-D Year operation began	0.0002		0.0003
γ_9	Title IV North Atlantic	2.79	**	0.890
γ_{10}	Title IV South Atlantic	0.496		0.932
γ_{11}	NSPS-D North Atlantic	-0.329		0.624
γ_{12}	NSPS-D South Atlantic	0.081		0.447
γ_{13}	NSPS-D South Central	-2.38	**	0.747
γ_{14}	NSPS-D Plateau	0.286		0.579
γ_{15}	Title IV dummy	-32.5	*	13.2
γ_{16}	Title IV ln(low switch cost)	0.959		0.735
γ_{17}	Title IV ln(high switch cost)	0.565		0.337
γ_{18}	NSPS-D ln(switch cost)	0.123		0.133
γ_{19}	Title IV ln(SO_2 difference)	1.19		0.927
γ_{20}	NSPS-D ln(SO_2 difference)	-0.164		0.320
γ_{21}	Minemouth	1.69	**	0.584
γ_{22}	Title IV 10-year coal contract	0.0004	*	0.0002
γ_{23}	Title IV Statutory bias	0.256		0.603
γ_{24}	Title IV Prior scrubber experience	0.777		0.599
γ_{25}	NSPS-D Prior scrubber experience	-2.14	**	0.414
γ_{26}	Cooperative	0.146		0.386
γ_{27}	Municipal	0.568		0.532
γ_{28}	State-owned	3.13	**	0.882
γ_{29}	Federal	1.38		2.19
γ_0	Constant	-8.06		5.31

Notes: * significant at 5 % level   ** at 1 % level

N = 405
Censored obs = 91
Log-likelihood = -101.1719
Wald test that all coefficients are equal to zero: $(\chi^2_{29}) = 657.05$

Table 6b: First-stage maximum likelihood estimates

Scrubbing cost equation

Dependent variable is ln(average cost of scrubbing)

Parameter	Variable	Coefficient		Standard error
β_1	Title IV ln(nameplate capacity)	0.201		0.333
β_2	NSPS-D ln(nameplate capacity)	0.027		0.293
β_3	Title IV ln(heat input)	-0.516		0.355
β_4	NSPS-D ln(heat input)	0.233		0.260
β_5	Title IV ln(SO_2 content)	-0.652		0.445
β_6	NSPS-D ln(SO_2 content)	-0.413	**	0.092
β_7	Title IV Boiler age in 1990	-0.013		0.016
β_8	NSPS-D Year operation began	-0.0002		0.0001
β_9	Title IV North Atlantic	0.035		0.253
β_{10}	Title IV South Atlantic	-0.449		0.250
β_{11}	NSPS-D North Atlantic	-0.277		0.239
β_{12}	NSPS-D South Atlantic	-0.043		0.136
β_{11}	NSPS-D South Central	-0.420	*	0.169
β_{12}	NSPS-D Plateau	0.709	**	0.142
β_{13}	Title IV dummy	8.64	*	3.75
β_0	constant	0.296		2.04

Notes: * significant at 5 % level ** at 1 % level

Censored N = 91

null hypothesis is that the coefficients on all of the unit characteristic variables in the log-likelihood function (12) are zero: that is, $H_0 : (\beta_1, \beta_2, \ldots, \beta_8, \gamma_1, \ldots, \gamma_8) = 0$. The likelihood function under the null is $\hat{L}_R = -147.709$; using the estimated unconstrained likelihood function given in Table 6a, the likelihood ratio test is $LR = 93.07$. The test statistic is distributed as a chi-squared variable with sixteen degrees of freedom; the critical value at a 1% significance level is 32. Similarly, likelihood ratio tests of the corresponding null hypotheses for each regime separately yield test statistics of 58.91 and 39.05 for Title IV and NSPS-D units, respectively, where each is distributed as a χ^2_7 with 1% critical value 18.48.[58] Hence for the model as a whole and for each regime considered separately, we can reject the hypothesis that the coefficients on the unit characteristics all equal zero.

Another measure of whether the importance of the regional dummies in determining scrubber costs affects the inferences from our model comes from the estimates of the structural equation, which also includes the regional dummies. As we show in the discussion of the structural technique choice equation, below, we can distinguish the effects of regional location from the effects of scrubbing and switching costs. This undermines the argument that the scrubber costs are driven only by regional location.

5.2.2 The structural technique choice equation

In the second stage, we use our estimates for $\hat{}$ to generate estimates of scrubber cost, $\hat{c}_i^s$. We can then estimate likelihood function (11) by the probit method. The results are provided in Table 7. These are our estimates of the structural technique choice equation – the heart of the model.[59] For ease of presentation, we have abused notation slightly and written these coefficients as θ and δ rather than θ/η and δ/η. The inferences discussed below are unaffected, since we focus on the signs and relative magnitudes of the estimated parameters. Note that Table 7 provides the negative of the coefficients on ln(scrubbing cost), θ_T^s and θ_N^s, to make interpretation easier. (Recall that scrubbing cost enters in negatively in the probability-of-scrubbing equation (10).)

Table 8 provides one measure of the "fit" of the model. We compare scrubber predictions with actual observations, where scrubber predictions are generated as a dummy variable that equals one if the estimated probability of scrubbing is greater than some cutoff value. Table 8 includes results for two natural cutoff values: 0.5 and the cutoff value that yields the same number of predicted scrubbers as actual scrubbers in the data. Using the latter cutoff, the model correctly predicts 83 out of 101 (82%) scrubbed units, and 296 out of 314 (94%) non-scrubbed units.

For convenience in interpreting the results, we present the estimates from the same model in slightly different form in Tables 9 and 10. In Table 9, the explanatory variables are the base variables (e.g., ln(scrubbing cost)) plus those variables interacted with a Title IV dummy. The new coefficients are marked with asterisks. Recall that Table 7 corresponds to the original equation, in which estimated scrubbing cost (for example) enters separately for Title IV and NSPS-D units. In the equation used to produce the results in Table 9, the corresponding variables are estimated scrubbing cost for all units and estimated scrubbing cost for Title IV units. Thus the coefficient $\hat{\theta}_T^{s*}$ in Table 9 is a measure of the difference between the coefficients for Title IV plants and NSPSD plants. Note that $\hat{\theta}^{s*} = \hat{\theta}_N^s$, and $\hat{\theta}_T^{s*} = \hat{\theta}_T^s - \hat{\theta}_N^s$. Similarly, $\hat{\theta}_T^{c*}$ is the "base" coefficient on switching

---

[58] The null hypotheses for the coefficients for Title IV and NSPS-D units, respectively, are $H_0 : (\beta_1, \beta_3, \beta_5, \gamma_1, \gamma_3, \gamma_5, \gamma_7) = 0$ and $H_0 : (\beta_2, \beta_4, \beta_6, \gamma_2, \gamma_4, \gamma_6, \gamma_8) = 0.$

[59] Note that the likelihood function (11) does not include a constant term, since the underlying probability-of-scrubbing equation given by (4) does not include one.

Table 7: Second-stage maximum likelihood estimates

Structural technique choice equation

Dependent variable is I_i = 1 for scrubbed units, I_i = 0 otherwise

Parameter[a]	Variable	Coefficient		Standard error
¡ μ^{s}_{T}	Title IV ln(scrubbing cost)	-1.36	*	0.555
¡ μ^{s}_{N}	NSPS-D ln(scrubbing cost)	0.251		0.135
μ^{c}_{TL}	Title IV ln(low switching cost)	1.18	**	0.331
μ^{c}_{TH}	Title IV ln(high switching cost)	0.519	*	0.242
μ^{c}_{N}	NSPS-D ln(switching cost)	0.272	*	0.117
μ^{d}_{T}	Title IV ln(difference in SO_2 content)	0.092		0.269
μ^{d}_{N}	NSPS-D ln(difference in SO_2 content)	-0.155		0.180
$\pm_1$	Minemouth	2.14	**	0.522
$\pm_2$	Title IV 10-year coal contract	0.0002	*	0.0001
$\pm_3$	Title IV Statutory bias	-0.512		0.351
$\pm_4$	Title IV Prior scrubber experience	0.239		0.313
$\pm_5$	NSPS-D Prior scrubber experience	-2.25	**	0.374
$\pm_6$	Title IV dummy	-1.56		1.29
$\pm_7$	Cooperative	0.223		0.342
$\pm_8$	Municipal	0.608		0.412
$\pm_9$	State-owned	3.42	**	0.808
$\pm_{10}$	Federal	-1.49	*	0.721
$\pm_{11}$	Title IV North Atlantic	0.924	*	0.388
$\pm_{12}$	Title IV South Atlantic	0.227		0.429
$\pm_{13}$	NSPS-D North Atlantic	-0.017		0.602
$\pm_{14}$	NSPS-D South Atlantic	0.023		0.384
$\pm_{15}$	NSPS-D South Central	-2.26	**	0.653
$\pm_{16}$	Plateau	-0.111		0.553

Notes: * significant at 5 % level ** at 1 % level

[a] As noted in the text, the estimates are actually of µ=´ and ±=´.
For the sake of presentation, we ignore the normalization by the standard error ´.

N = 415
Number of units with scrubbers = 101
Log-likelihood = ¡ 103:02

Table 8: Predicted vs. actual scrubbing choices

Cutoff = 0.5

Actual	Predicted 0	1	Total
0	312	2	314
1	28	73	101
Total	340	75	415

Cutoff = 0.281

Actual	Predicted 0	1	Total
0	296	18	314
1	18	83	101
Total	314	101	415

costs for Title IV plants, while $\hat{\theta}^{c*}_{TL}$ measures the difference between the coefficients on low and high switching costs.

Table 10 presents the probit results in terms of the marginal impact of each variable on the probability of scrubbing. These marginal effects for each variable are measured separately for units under each regime, at the means of the other variables. Note that the results for the continuous variables are expressed in terms of the units of the original (not logarithmic) values of the variables.[60] The effects for the dummy variables correspond to a discrete change in the variable from zero to one. For the ownership and regional dummies, the effects are measured as the difference in estimated scrubbing probability relative to an investor-owned or North Central unit.

Comparing cost sensitivity across policy regimes First, consider the evidence on the scrubbing decisions of Title IV units. Our estimate of $-\theta^s_T$ is negative and significant at the 5% level. This suggests that Title IV units are indeed sensitive to scrubbing costs: the probability of scrubbing is less, the higher is the estimated cost of scrubbing. From Table 10, an increase in scrubbing cost of one cent per pound SO_2 corresponds to a 2.5% drop in the probability of scrubbing.

Our estimates of the coefficients on coal-switching cost, $-\hat{\theta}^c_{TL}$ and $\hat{\theta}^c_{TH}$, are also positive and significantly different from zero (at the 1% and 5% levels respectively): greater switching costs make scrubbing more likely. An increase of one cent per million Btus in the cost of coal-switching

---

[60] The marginal effect of the untransformed variables in the scrubbing cost equation is given by

$$\begin{aligned}\frac{\partial \Pr(\text{scrub})}{\partial X_i} &= \frac{\partial \Pr(\text{scrub})}{(\partial X_i/X_i)X_i} \\ &= \frac{\partial \Pr(\text{scrub})}{\partial x_i} \cdot \frac{1}{X_i} \\ &= \frac{\phi(x_i\beta)\beta}{X_i};\end{aligned}$$

where as before $x_i = \ln X_i$.

Table 9: Second-stage maximum likelihood estimates
Coefficients for Title IV units expressed
relative to NSPS-D units
Dependent variable is $I_i = 1$ for scrubbed units, $I_i = 0$ otherwise

Parameter[a]	Variable	Coefficient		Standard error
$-\theta^{s*}$	ln(scrubbing cost)	0.251		0.135
$-\theta^{s*}_T$	Title IV ln(scrubbing cost)	-1.61	**	0.574
θ^{c*}	ln(switching cost)	0.272	*	0.118
θ^{c*}_T	Title IV ln(switching cost)	0.903	**	0.345
θ^{c*}_{TH}	Title IV ln(high switching cost)	-0.656	**	0.165

Notes: * significant at 5 % level ** at 1 % level

is correlated with an increase of 6.5% in the probability of scrubbing. The effect of high switching costs is much smaller.

The relative magnitudes of the estimated coefficients for Title IV units also accord with theory. The managers of Title IV units in general react similarly to scrubbing and switching costs: the estimates of $-\theta^s_T$ and $-\theta^c_{TL}$ are not significantly different from one another. Meanwhile, note that $\hat{\theta}^{c*}_{TL}$ in Table 9 is positive and significantly different from zero at the 5% level. Hence Title IV units respond more to changes in switching costs when the cost of switching is less than the allowance price, as expected.

The estimated effect of scrubbing costs on the probability of scrubbing by NSPS-D units, $\hat{\theta}^c_N$, is not significantly different from zero at the 5% level. Thus we cannot reject the hypothesis that NSPS-D plants ignore scrubbing costs completely in their scrubbing decision.[61] We do, however, find a significant response among NSPS-D units to the cost of switching: $\hat{\theta}^c_N$ is positive and significantly different from zero at the 5% level. The magnitude of the response is quite small: the probability of scrubbing increases by about one-half of one percent when estimated switching costs increase by a cent per million Btus.

A comparison between the two regimes strongly supports the hypothesis that Title IV units are more responsive to average abatement costs than are NSPS-D plants. From Table 9, $\hat{\theta}^{s*}_T$ is positive and significant at the 1% level, implying that an increase in scrubbing cost has a greater effect on the probability of scrubbing for Title IV units than for NSPS-D units. Our estimate of $\hat{\theta}^{c*}_T$ is also positive and significant at the 1% level, implying that Title IV units are more responsive to the cost of coal-switching as well.

Other influences on the choice of technique Table 7 also provides evidence for other influences on the scrubbing decision.[62] The estimated coefficient on minemouth location, $\hat{\delta}_1$, is positive,

---

[61] Indeed, if anything, the evidence seems to indicate that units with higher scrubbing costs are somewhat more likely to choose scrubbing.

[62] Neither of the coefficients on the difference in coal content is significantly different from zero. Recall that these variables are included to correct for the fact that different units face different high- and low-sulfur coals. The

Table 10: Second-stage maximum likelihood estimates
Estimates of marginal effects of variables
on the probability of scrubbing

Dependent variable is $I_i = 1$ for scrubbed units, $I_i = 0$ otherwise

Variable	Units	Marginal effect on Prob(scrubbing)[a]
Title IV units		
Scrubbing cost	cents/lb SO_2	-0.025*
Low switching cost	cents/mmBtu[b]	0.065**
High switching cost	cents/mmBtu	0.002*
Diff. in SO_2 content	lbs SO_2/mmBtu	0.003
Minemouth		0.642**
Prior scrubber experience		0.025
Ten-year contract[c]	000 tons	1.8 x 10e-5*
Statutory bias		-0.045
Cooperative[d]		0.033
Municipal		0.116
Federal		-0.062*
North Atlantic[e]		0.135*
South Atlantic		0.018
NSPS-D units		
Scrubbing cost	cents/lb SO_2	0.004
Switching cost	cents/mmBtus	0.005*
Diff. in SO_2 content	lbs SO_2/mmBtu	-0.026
Minemouth		0.677**
Prior scrubber experience		-0.738**
Cooperative[d]		0.073
Municipal		0.219
State-owned		0.763**
North Atlantic[e]		-0.007
South Atlantic		0.009
South Central		-0.582**
Plateau		-0.043

Notes: * significant at 5 % level ** at 1 % level
where significance levels are taken from Table 9.

[a] With one exception, entries are $(\partial\hat{\Phi}/\partial x_i)/X_i$ for continuous variables and $\hat{\Phi}(d_i = 1) - \hat{\Phi}(d_i = 0)$ for dummy variables, where $\hat{\Phi}$ is the estimated probability of scrubbing.

[b] Note that the units for scrubbing and switching cost are different. The correction comes by including the difference in SO_2 content, which is measured in lbs SO_2/mmBtu.

[c] Entry is simply $(\partial\hat{\Phi}/\partial x_i)$, since variable is already in level form.

[d] Effects of ownership dummies are relative to investor ownership.

[e] Effects of regional dummies are relative to the North Central region.

significantly different from zero at the 1% level, and large. It implies that minemouth units are much more likely to install scrubbers, in line with our expectations. From Table 10, a minemouth unit under either policy regime is roughly two-thirds more likely to install a scrubber than an otherwise identical unit located away from a mine.

The probability that a unit installs a scrubber also increases with the amount of coal under long-term contract, consistent with expectations. Our estimate of $\hat{\delta}_2$ is positive and significantly different from zero at the 5% level. The effect is small, but not negligible. An increase at the plant level of roughly 500,000 tons of coal under long-term contract raises the probability of scrubbing by one percent.[63]

Neither statutory bias or prior scrubbing experience for Title IV units has a significant estimated effect.[64] However, our estimate for the coefficient on prior experience among NSPS-D units, $\hat{\delta}_5$, is negative, large in absolute value, and significant at the 1% level. This result might suggest a "negative learning" process, in which utilities using scrubbers are disappointed with their performance and thus less likely to install them again. Alternatively, because prior scrubbing experience is more likely for boilers built later, the negative coefficient might reflect a fall in the probability of scrubbing over time.[65]

Note that the estimated coefficient on the dummy variable for Title IV units, $\hat{\delta}_6$, is negative, but not significantly different from zero. The sign is consistent with an option value from waiting, although – perhaps not surprisingly – we are unable to distinguish the effect from zero.

The coefficients on the ownership dummies indicate that public ownership, at least by state or federal entities, has an effect on technique choice. The estimated coefficient on the state-ownership dummy, $\hat{\delta}_9$, is positive and significant at the 1% level, suggesting that state-owned units may be under greater pressure to use scrubbers rather than switch coal. From Table 10, this effect is economically significant as well: compared to similar investor-owned units, state-owned units are 76% more likely to have scrubbers. Such pressure might come either from in-state coal mining interests or from environmentalists supporting scrubbing because it leads to lower emissions of sulfur dioxide. On the other hand, the coefficient on the federal-ownership dummy, $\hat{\delta}_{10}$; is negative and significant at the 5% level, suggesting that federally owned units are less likely to install scrubbers than other units.[66] The estimated effect, around 6%, is much smaller than that of state ownership.

Finally, two coefficients on the regional dummies are significantly different from zero. Title IV units in the North Atlantic region are 13.5% more likely to have scrubbers than similar units in the North Central region, all else equal; this effect is significant at the 5% level. On the other

---

coefficients on these variables are somewhat hard to interpret. Because the "correction" mentioned involves dividing by the difference in SO_2 content, we might expect the sign of the coefficient to be negative. On the other hand, the difference in SO_2 content is clearly correlated with the actual SO_2 content of coal burned. If units burning higher sulfur coal are more likely to adopt scrubbers, all else equal, then the coefficient on ln(difference in SO_2 content) might be biased upward: higher SO_2 content increases both the probability of adoption and the difference in SO_2 content.

[63] By comparison, average coal consumption at the unit level for Title IV units was 827,000 tons in 1998.

[64] The lack of evidence that state regulatory biases encouraged scrubbing is consistent with the conclusion of Ellerman et al. (2000) that state regulations do not appear to have discourage allowance trading.

[65] However, because we include time trends in the equations used to estimate costs, such a change would presumably not be due to changes in relative costs over time.

[66] Of course, in our sample federal ownership is synonymous with units operated by the Tennessee Valley Authority; thus the "federal" ownership dummy is essentially a fixed effect for the TVA. Still, it is not obvious why units at TVA plants would be less likely to scrub than investor-owned units.

Table 11: Second-stage maximum likelihood estimates
Structural technique choice equation
Alternative regression, without regional dummy variables
Dependent variable is $I_i = 1$ for scrubbed units, $I_i = 0$ otherwise

Parameter[a]	Variable	Coefficient		Standard error
$-\theta^s_T$	Title IV ln(scrubbing cost)	-0.939	*	0.440
$-\theta^s_N$	NSPS-D ln(scrubbing cost)	0.108		0.088
θ^c_{TL}	Title IV ln(low switching cost)	1.20	**	0.339
θ^c_{TH}	Title IV ln(high switching cost)	0.596	*	0.253
θ^c_N	NSPS-D ln(switching cost)	0.247	*	0.095
θ^d_T	Title IV ln(difference in SO_2 content)	0.075		0.225
θ^d_N	NSPS-D ln(difference in SO_2 content)	-0.088		0.183
δ_1	Minemouth	2.29	**	0.500
δ_2	Title IV 10-year coal contract	0.0002	**	0.0001
δ_3	Title IV Statutory bias	-0.609		0.320
δ_4	Title IV Prior scrubber experience	0.145		0.296
δ_5	NSPS-D Prior scrubber experience	-1.97	**	0.290
δ_6	Title IV dummy	-2.05		1.12
δ_7	Cooperative	0.206		0.300
δ_8	Municipal	0.518		0.369
δ_9	State-owned	1.81	**	0.497
δ_{10}	Federal	-1.20	*	0.606

Notes: * significant at 5 % level ** at 1 % level

[a] As noted in the text, the estimates are actually of θ/σ and δ/σ.
For the sake of presentation, we ignore the normalization by the standard error σ.

$N = 415$
Number of units with scrubbers = 101
Log-likelihood = -117.63

hand, a NSPS-D unit in the South Central region is 58% less likely to have a scrubber than an otherwise similar unit in the North Central region, at a 1% significance level. These results are broadly consistent with the effects of state environmental regulations, which are strongest in the Northeast and the Atlantic coast regions (recall that California is not represented in our sample) and weakest in states such as Louisiana, Arkansas, and Texas, all in the South Central region.

Of course, because dummy variables act as "catch-all" variables, interpreting their precise meaning is difficult. Because of their apparent significance in the regression, it is also worthwhile assessing the robustness of the other results to the inclusion of regional dummies. Table 11 presents results from estimating the second-stage structural equation without the regional dummy variables.

Note two key results from the alternate regression. First, the regional dummies contribute significant explanatory power to the model. The LR test statistic for the null hypothesis that all the dummies equal zero, given by $H_0 : (\delta_{10}, \ldots, \delta_{15}) = 0$, is 32.9; the critical value for the χ^2_6

distribution at the 1% level is 16.81.

Second, the regional dummies have only a minor impact on the estimated values of the other coefficients. The magnitudes of the estimates change, but – except for the coefficients on a few variables – the changes are within one standard error. None of our earlier inferences would be reversed, although the estimated coefficient on statutory bias (δ_3) becomes significant at the 5% level.[67] This appears to be due to omitted variable bias. The former result is most notable, because it suggests that NSPS-D units are more likely to choose scrubbing as the costs of scrubbing increase.

5.2.3 A note on the use of generated regressors

In estimating the probability of scrubbing, we use econometrically generated regressors: estimates of scrubbing and switching cost, rather than direct observations. This potentially introduces a specific kind of errors-in-variables problem, in which we have estimates of the size of the error – namely, the forecast errors from the original regressions used to generate the cost estimates.[68] A general treatment of this problem is provided by Murphy and Topel (1985) [MT]. Two results are particularly important here. First, the ML parameter estimates are consistent. This result (which does not hold in the general "errors in variables" problem) follows because the estimated regressors are themselves consistent estimates of the underlying variables. That is, on the assumption that our models of switching costs and scrubbing costs are correctly specified, the fact that we estimate those costs rather than observe them directly does not bias our parameter estimates in the later-stage models.

The standard errors, however, are affected by the use of econometric estimates, since the usual calculation of the variance matrix does not take into account the fact that the regressors are estimated. In general, the direction of the bias (i.e., whether the corrected standard errors would be smaller or larger) cannot be determined without calculating the correction.[69] On the other hand, in the specific case in which the disturbance term in the model used to estimate the regressor is independent from the disturbance term in the later-stage model, MT show that the uncorrected standard errors underestimate the true standard errors. This case appears to apply to the switching cost estimates: it is reasonable to assume that the disturbance term in the coal-price regressions is independent of the disturbance term in the scrubbing equation. Thus we would expect the corrections to result in larger standard errors. Although this puts some of our inferences in doubt, many of our results above are significant at the 1% level, and would likely be robust to such a correction.

It is also worth pointing out that the measurement error "problem" we have identified may not be a problem at all. As has frequently been observed, the measurement error problem is only a problem to the extent that what the econometrician measures is different from what the economic

---

[67] Another effect is the substantial fall in the estimated coefficient on state ownership, $\hat{\delta}_8$. This can be explained by omitted-variable bias in the alternate regression. State ownership is most prevalent in the South Central region, where scrubbing is relatively uncommon. Omitting the regional dummies thus attributes the regional effect to state ownership.

[68] Estimation using an instrumental variables approach does not seem suitable, since we lack any observation on the underlying independent variable of interest (scrubbing or switching cost).

[69] While MT provide methods for correcting standard errors when the regressors have been econometrically estimated, our model poses a greater challenge, because we have (at least in the structural technique choice equation) two such generated regressors, produced by entirely different earlier-stage equations. Partly for that reason, we have not yet attempted the corrections. However, we plan on doing so in future versions of this paper.

agents react to. In our case, we are interested in the sensitivity of the choice of technique to the costs of switching and scrubbing. It is fair to ask what information managers of electric power plants are likely to base their decisions on. To the extent that managers get accurate and site-specific estimates of what a scrubber would cost, or to the extent they have detailed knowledge of what price they would have to pay for other types of coal, then our estimates necessarily introduce measurement error. On the other hand, one might imagine that managers base their decisions on estimates of average scrubbing costs for "units like theirs," or on information gleaned about prevailing coal price patterns at other plants nearby. In that case, our measurements may be very good proxies for the information available to the manager, and the measurement error problem may be moot. Indeed, one might argue that sophisticated managers would do exactly the kind of exercise we have performed here – that is, gather data and run regressions – in order to generate estimates of scrubbing and switching costs.

6 Conclusion

We have analyzed the determinants of pollution abatement technique choice, focusing in particular on the effects of environmental policy instruments. The model developed in Section 2 shows that cost savings from new technologies should be greater under a market-based instrument than under an equivalent command-and-control policy. Hence the model predicts that the technique choice of firms under a system of tradeable permits will be more sensitive to abatement costs than the choice of firms under an emissions rate standard.

We tested this prediction using data from coal-fired electric generating units. The results presented here provide empirical evidence that a market-based instrument provides greater incentives for the adoption of new pollution control technologies than a uniform emissions rate standard. The technique choice decisions of managers – to scrub or to switch coals – were responsive to changes in the costs of scrubbing or coal-switching for units under the Title IV regime. Increases in scrubbing cost made scrubbing less likely, while increases in the costs of coal switching had the opposite effect. Moreover, these units were more sensitive to changes in costs than were units under the NSPS-D regime: a given change in switching cost, for example, was associated with a greater increase in the probability of scrubbing.

Several extensions are possible. The empirical analysis could be extended to simulate the effects of the policy regime on patterns of technique choice – for example, one could use the choice model estimated here to predict the counterfactual distribution of scrubbers made if an emissions rate standard had been used, say, instead of Title IV. Connected to a model of health and environmental effects of sulfur dioxide, such simulation results could shed light on the indirect environmental effects of choosing one policy instrument versus another.

On the theoretical front, the simple model developed here provides a foundation for a more general model of environmental policy and technology adoption, taking into account heterogeneity among polluters; output decisions by firms; industry structure in the output market and among the suppliers of abatement technology; and so on. A more complete model would also compare the various policy instruments in terms of their welfare effects.

The work on generating price premia for low-sulfur coal could also be fruitfully extended in a number of directions. In particular, no theoretical model of price premia exists, although the phenomenon is one we can expect to find in any situation in which firms can comply with environmental regulations by changing fuels.[70] The market for low-sulfur coal after 1995 also represents an interesting case of a nascent market; using the price premia developed here, one could examine the extent to which sulfur premia and prices of sulfur dioxide allowances converged over the first few years of Title IV.

For emerging issues such as global climate change, the time horizon is measured in decades rather than years. An understanding of how policy instruments affect the evolution of pollution control technology is crucial for designing environmental policy for the long term. Analyzing technique choice under different environmental policy instruments is a first step towards such a broader understanding.

---

[70] For example, the imposition of a tax on carbon content, or the institution of a system of tradeable carbon rights, would be expected to generate a premium on lower-carbon fuels.

Appendix: Estimation of coal prices

For each coal district listed in Table 3, points of origin by barge and by each railroad line were chosen. Origins were chosen using three pieces of information. Where possible, the origin used by the weekly trade publication "Coal Outlook" was used.[71] For example, rail distances for Central Appalachian coal to points west, north, and northeast were measured from Kenova (for the Norfolk Southern railroad) and Louisa, on the banks of the Big Sandy River (for the CSX railroad). Likewise, barge distances for Central Appalachian coal were measured from the mouth of the Big Sandy River. Where "Coal Outlook" failed to name a precise point of origin, we relied on the geographical concentration of coal mining, as depicted by the Coal Activity Map, along with the proportion of coal originating from different counties in the district (as given by the data from Form 423). Where appropriate, different origins were chosen for different destination regions. For example, we assumed that Central Appalachian coal bound for Georgia would be shipped south from Central Appalachia, along rail connections leading through Tennessee, rather than north through Kenova/Louisa in West Virginia.

For transport by barge, in cases where the coal district was not adjacent to water (or for transport to other water bodies), ports were identified for each coal district. For example, ports were identified for transport of Central Appalachian coal to barge-served plants on the Great Lakes. Similarly, ports were identified or transport of high-sulfur coal from western Kentucky up and down the Mississippi. For western coal – in particular, coal from the Powder River Basin in Wyoming – ports were chosen along the Mississippi and for the Great Lakes. These ports were chosen by first identifying the closest major ports (by rail) to each coal district in question, for various bodies of water. Then, we chose among "candidate ports" by referring to figures for coal shipments given in the U.S. Army Corps of Engineers' reports on "Waterborne Commerce of the United States" for 1998 (USACE 1998). For example, examination of shipments of coal on the Mississippi River reveals that St. Louis is the major port, far outstripping Minneapolis (for example); thus St. Louis was used as the port for western coal delivered to barge-served plants on the Mississippi.

Railroad distances to each rail-served plant were then calculated from the railroad atlas. Wherever possible, railroad mileages were calculated using the actual railroad serving a plant. Distances were measured to the nearest town to a power plant, as provided by Form 767. In instances where rail distances could not be calculated directly from the railroad atlas (e.g., the nearest town did not appear in the railroad atlas), the road atlas was used to estimate distances from the nearest railroad junction. Barge distances from origins or ports were measured on the state-level road atlas, along with calculations using river mileage posts provided on the "Coal Activity Map."

For each district, by barge and by rail, real spot market[72] coal prices (in 1996 cents per million Btus) were then regressed on several variables: distance to the coal district; ash, sulfur, and heat content of the coal; a dummy variable for compliance coal (coal that when burned would produce less than 1.2 pounds of SO_2 per mmBtu, thus allowing compliance with the NSPS regulations without scrubbing), for appropriate districts; nameplate capacity of the plant; a dummy for plants

---

[71] Sample issues of "Coal Outlook" are available online at http://www.ftenergyusa.com/coaloutlook/default.asp.

[72] We elected to use spot market data, rather than data on coal deliveries under contract, because the spot delivery data is likely to be a better representation of the prices prevailing in the market at a given point in time. Hence it provides a better foundation for predicting counterfactual coal prices. As long as the price differentials for spot-market coal from various districts follow the same pattern as price differentials in coal contracts, price premia derived from estimated spot market prices will provide reasonable proxies for the "true" price premia that would face individual plants.

that were served by both barge and rail; annual dummies; dummies for the railroad or water body serving the plant; dummies (where appropriate) for the port[73]; and, where the data permitted, interactions between distance and annual dummies and between distance and railroad or water body dummies. The fits of the regressions were generally good: R^2 typically fell in the 0.70-0.80 range.[74] The regressions on PRB coal had a slightly worse fit, with R^2s between 0.60 and 0.70.

Tables A1 and A2 provide two sample price regressions: for Central Appalachian coal by barge and PRB coal by rail, over the years 1985-1998.[75] These are straightforward OLS regressions. The observations are individual coal deliveries reported in the Form 423 data, for plants for which we have distance data.[76] A few of the variables { sulfur, ash, and heat content, and of course the price { pertain to the coal delivered. Most of the other variables are characteristics of the plant receiving the coal. Note that the regression for barge-served plants has a dummy for rail-served plants, and vice versa; some plants are served by both modes of transport. In the barge regression, we have included dummies { some interacted with distance { for the water bodies of receiving plants (e.g., Atlantic, Great Lakes, Mississippi; in the Central Appalachian regression the major omitted dummy is for the Ohio River system). Dummies for ports (e.g., Chicago, Toledo) are also included. In the rail regression, we have dummy variables for railroad lines: here, BNSF and UP. Variables followed by a year are year dummies (e.g., \Year 1985") or other variables interacted with the corresponding year dummies (e.g., \Distance 1985").

These regressions were then used to generate estimated prices for coal from a variety of districts in every year from 1985-1998 for each plant. Prices for coal from the two major districts, Central Appalachia and the Powder River Basin, were generated for every plant. The other districts chosen for each plant depended on the plant's location, and on the districts it had purchased coal from. Prices for coal from district 13 (Alabama) were calculated for plants near Birmingham, but not Baltimore.

---

[73] Note that the inclusion of dummy variables for the ports compensates for the unavoidable uncertainty in our selection of the ports.

[74] The reader may note that the regressions in Tables A1 and A2 do not include a constant term. Values for R^2 were calculated from the equivalent regressions with one dummy variable replaced by a constant term.

[75] The full set of regression results, along with details on the variables and the methods used, can be found in [paper title TBA { work in progress on coal price premia].

[76] We have distance data for almost all of the plants of 100MW or greater capacity, not only those under the Title IV and NSPS-D regimes.

Table A1: Coal price regression
Central Appalachian coal by barge

Variable	Coefficient	Standard Error
Sulfur	-4.77	0.407
Sulfur 1990s	3.11	0.483
Sulfur 1995-98	-5.56	0.525
Compliance coal	3.28	0.236
Nameplate capacity	$-1.27\text{x}10^{-6}$	$1.41\text{x}10^{-7}$
Ash content	-1.23	0.053
Heat content	0.003	$2\text{x}10^{-4}$
Rail-served plant	-3.96	0.234
Atlantic	47.0	1.39
Great Lakes	29.4	0.710
Gulf	54.8	9.09
Mississippi	16.5	0.61
Tennessee	-28.4	2.61
Chicago	18.6	2.73
Chicago 1990-94	-6.36[a]	3.18
Chicago 1995-98	-10.8	3.83
Toledo 1980s	9.94	0.723
Jersey City 1990s	8.10	1.53
Baltimore	8.95	1.69
Baltimore 1990s	-9.78	1.88
Distance 1985	0.028	0.001
Distance 1986	0.025	0.001
Distance 1987	0.026	0.001
Distance 1988	0.022	0.001
Distance 1989	0.019	0.001
Distance 1990	0.020	0.001
Distance 1991	0.022	0.001
Distance 1992	0.022	0.001
Distance 1993	0.017	0.001
Distance 1994	0.020	0.001
Distance 1995	0.021	0.001
Distance 1996	0.022	0.001
Distance 1997	0.027	0.001
Distance 1998	0.025	0.001
Gulf distance	-0.022	0.005
Tenn. distance	0.032	0.003
Gr. Lakes dist.	0.008	0.003
Year 1985	134	3.80
Year 1986	126	3.82
Year 1987	114	3.82
Year 1988	105	3.83
Year 1989	108	3.81
Year 1990	108	3.80
Year 1991	95.6	3.80
Year 1992	92.0	3.82
Year 1993	102	3.83
Year 1994	99.9	3.81
Year 1995	90.4	3.81
Year 1996	90.6	3.79
Year 1997	90.5	3.80
Year 1998	91.5	3.80

Notes: All variables significant at 1% level, except for [a] at 5% level.
See text for explanation of variables.
$N = 21382$ true $R^2 = 0{:}81$

Table A2: Coal price regression
Powder River Basin coal by rail

Variable	Coefficient	Standard Error
Sulfur	49.3	11.4
Sulfur 1990s	-28.8	12.6
Sulfur 1995-98	-21.0	7.51
Ash content	-1.96	0.280
Heat content	0.013	0.001
Barge-served plant	-8.33	0.923
Other-served plant[a]	34.3[b]	13.9
Minemouth plant	50.0[b]	19.6
East coast plant	25.3	1.53
BNSF railroad	-19.9	2.69
UP railroad	-4.73	0.929
Distance 1985	0.067	0.006
Distance 1986	0.067	0.008
Distance 1987	0.044	0.005
Distance 1988	0.035	0.005
Distance 1989	0.055	0.004
Distance 1990	0.046	0.003
Distance 1991	0.045	0.003
Distance 1992	0.042	0.003
Distance 1993	0.048	0.003
Distance 1994	0.044	0.003
Distance 1995	0.039	0.003
Distance 1996	0.033	0.003
Distance 1997	0.037	0.003
Distance 1998	0.043	0.003
BNSF distance	0.026	0.002
Year 1985	-23.0[c]	13.6
Year 1986	-39.2	15.1
Year 1987	-28.1[b]	12.9
Year 1988	-32.2[b]	13.1
Year 1989	-58.0	12.4
Year 1990	-45.0	11.7
Year 1991	-46.0	11.7
Year 1992	-47.0	11.7
Year 1993	-54.0	11.7
Year 1994	-50.1	11.4
Year 1995	-45.2	11.5
Year 1996	-37.6	11.6
Year 1997	-42.0	11.7
Year 1998	-48.8	11.6

Notes: All variables significant at 1% level, except for [b] – at 5% level and [c] – at 10% level.

See text for explanation of variables.

N = 5287 true $R^2 = 0.64$

References

[1] Baumol, William J. and Wallace E. Oates. 1988. The Theory of Environmental Policy. Cambridge: Cambridge University Press.

[2] Bellas, Allen S. 1998. Empirical evidence of advances in scrubber technology. Resource and Energy Economics 20: 327-343.

[3] Downing, P. B. and L. J. White. 1986. "Innovation in pollution control." Journal of Environmental Economics and Management 13: 18-29.

[4] Energy Information Administration (EIA). 1985-1998. "Steam-Electric Plant Operation and Design Report." Form EIA-767.

[5] Ellerman, A. Denny, Paul L. Joskow, Richard Schmalensee, Juan-Pablo Montero, and Elizabeth M. Bailey. 2000. Markets for Clean Air: The U.S. Acid Rain Program. Cambridge: Cambridge University Press.

[6] Federal Energy Regulatory Commission (FERC). 1972-1998. "Monthly Report of Cost and Quality of Fuels for Electric Power Plants." FERC Form 423.

[7] Fischer, Carolyn, Ian W. H. Parry, and William A. Pizer. 1998. "Instrument choice for environmental protection when technological change is endogenous." RFF Discussion Paper 99-04 (October). Washington, D.C: Resources for the Future.

[8] Heckman, James J. 1974. "Shadow prices, market wages, and labor supply." Econometrica 42(4): 679-694.

[9] Heckman, James J. 1976. "The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models." Annals of Economic and Social Measurement 5/4: 475-492.

[10] Jaffe, Adam B. and Robert N. Stavins. 1995. "Dynamic incentives of environmental regulations: The effects of alternative policy instruments on technological diffusion." Journal of Environmental Economics and Management 29: S-43–S-63.

[11] Jaffe, Adam B. Richard G. Newell, and Robert N. Stavins. "Technological change and the environment." Forthcoming in The Handbook of Environmental Economics, Karl-Goran Maler and Jeffrey Vincent, eds. Amsterdam: North-Holland/Elsevier Science.

[12] Jung, C., K. Krutilla, and R. Boyd. 1996. "Incentives for advanced pollution abatement technology at the industry level: An evaluation of policy alternatives." Journal of Environmental Economics and Management 30: 95-111.

[13] Keohane, Nathaniel O. 1999. "Policy instruments and the diffusion of pollution abatement technology." Mimeo, Harvard University.

[14] Lee, Lung-Fei. 1976. "Unionism and wage rates: A simultaneous equations model with qualitative and limited dependent variables." International Economic Review 19(2): 415-433.

[15] Lile, Ron, and Dallas Burtraw. 1998. "State-level policies and regulatory guidance for compliance in the early years of the SO_2 emission allowance trading program." RFF Discussion Paper 98-35 (May). Washington, D.C.: Resources for the Future.

[16] Maddala, G. S. 1983. Limited-Dependent and Qualitative Variables in Economics. Cambridge: Cambridge University Press.

[17] Malueg, D. A. 1989. "Emission credit trading and the incentive to adopt new pollution abatement technology." Journal of Environmental Economics and Management 16: 52-57.

[18] Murphy, Kevin M., and Robert H. Topel. 1985. "Estimation and inference in two-step econometric models." Journal of Business and Economic Statistics 3(4): 370-379.

[19] Newell, Richard G., Adam B. Jaffe, and Robert N. Stavins. 1999. "The induced innovation hypothesis and energy-saving technological change." Quarterly Journal of Economics 114: 941-975.

[20] Rand McNally. 1982. Rand McNally Handy Railroad Atlas of the United States.

[21] Rand McNally. 1994. Rand McNally United States Road Atlas.

[22] Resource Data International. 1998. 1998 U. S. Coal Activity Map. Wall-size map.

[23] Schumpeter, Josef A. 1939. Business Cycles. New York: McGraw-Hill.

[24] United States Army Corps of Engineers (USACE). 1998. Waterborne Commerce of the United States, parts 1-3. Fort Belvoir, VA: Water Resources Support Center, U. S. Army Corps of Engineers.

[25] United States Census Bureau. 1972-1999. Statistical Abstract of the United States. Washington, D.C: xxx.

[26] Yatchew, Adonis, and Zvi Griliches. 1984. "Specification error in probit models." The Review of Economics and Statistics 66: 134-139.

FIGURE 1: Cost functions of three abatement techniques.

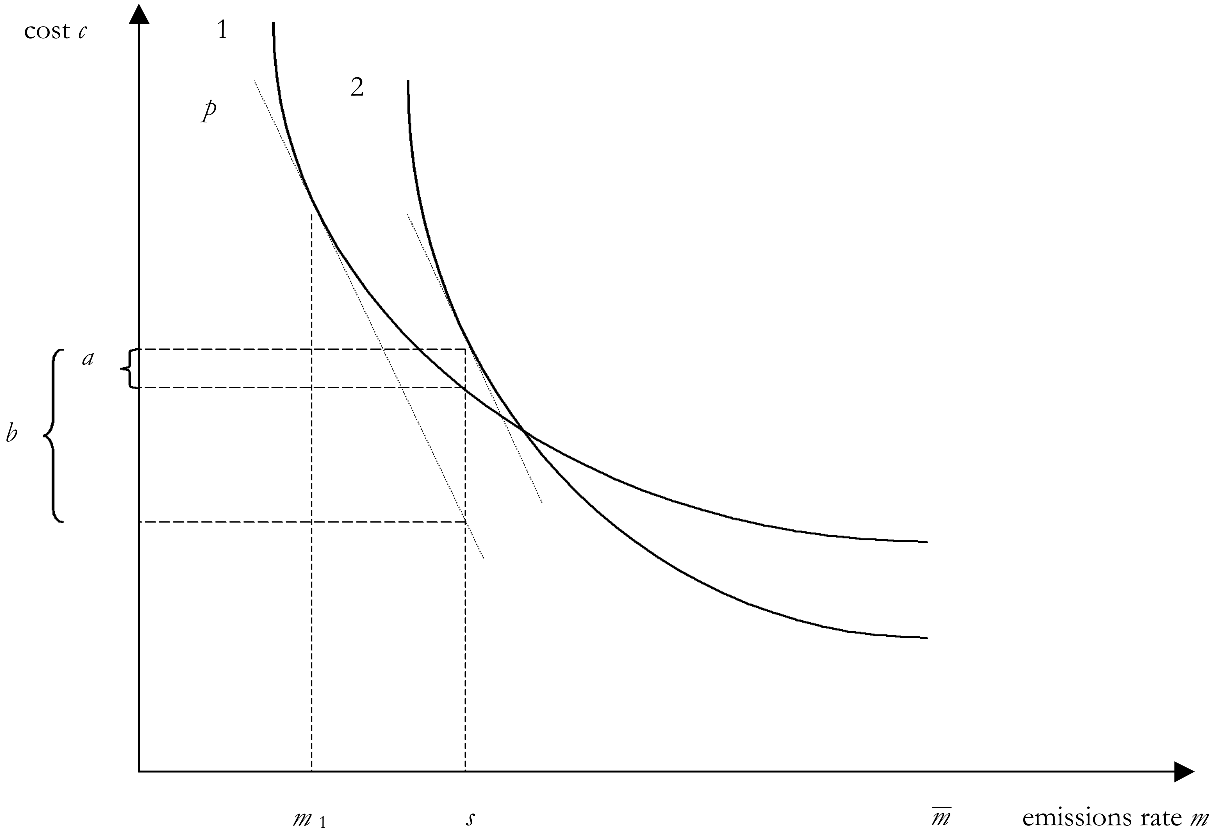

FIGURE 2A: The cost savings from the new technique (1) are greater under the market-based instrument than under the standard

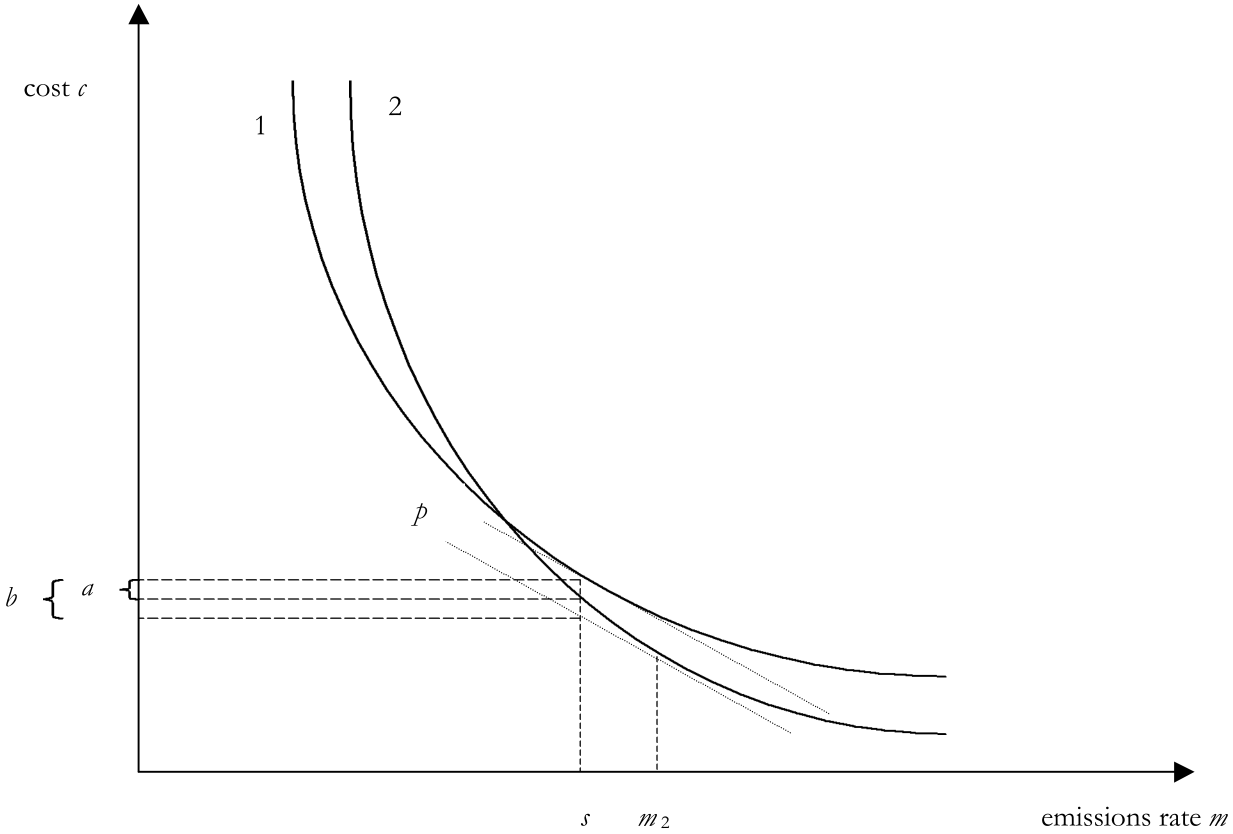

FIGURE 2B: New technique is technique (2).

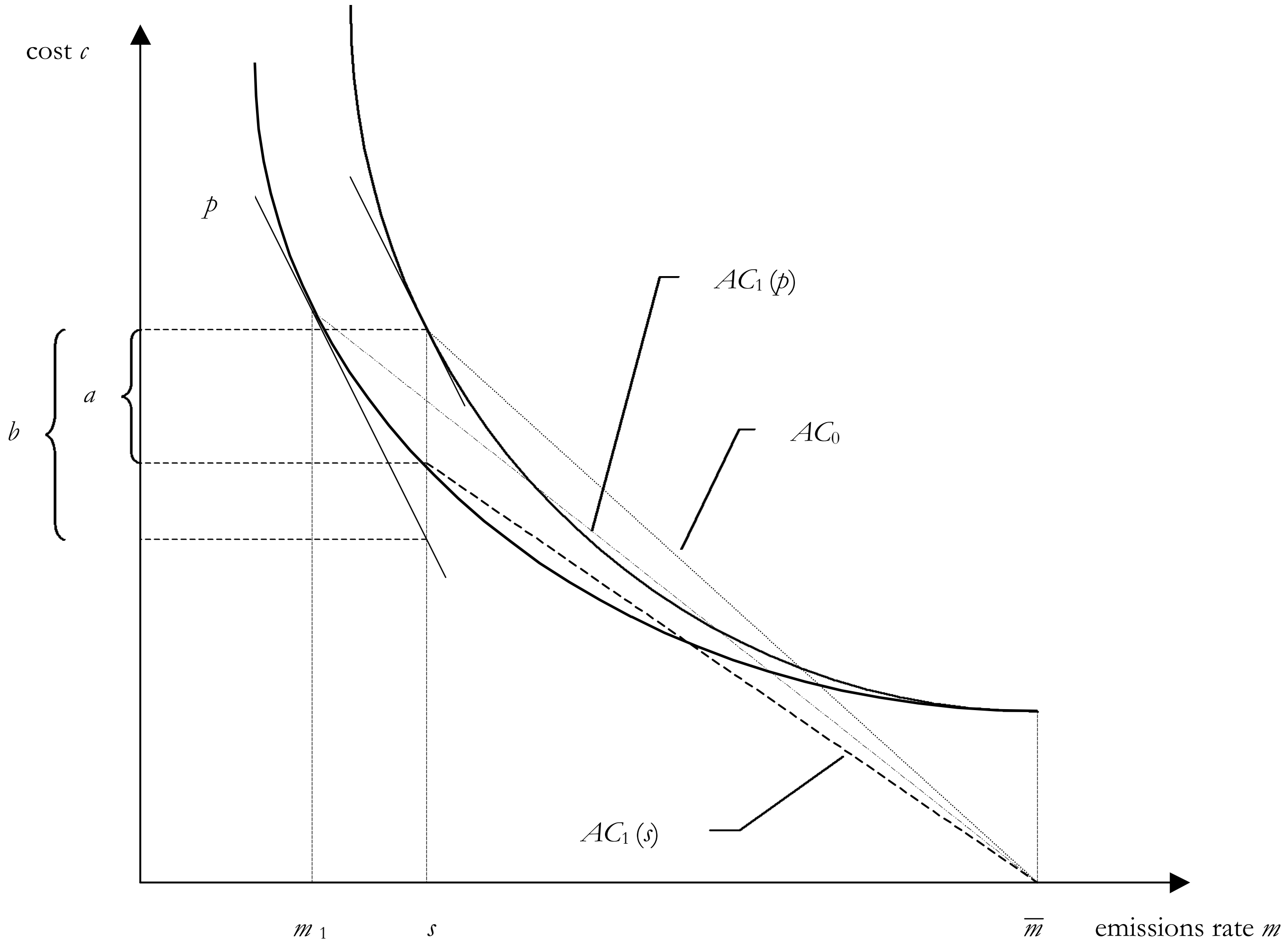

FIGURE 3: A given change in the cost function produces a larger change in total cost under the market-based instrument (*b*) than under the standard (*a*).