5-2010

Plant vintages, grandfathering, and environmental policy

Garth Heutel
Georgia State University, gheutel@gsu.edu

Plant Vintages, Grandfathering, and Environmental Policy

Garth Heutel
Department of Economics, University of North Carolina Greensboro
PO Box 26165
Greensboro, NC 27402
gaheutel@uncg.edu
336-334-4872 (Phone)
336-334-5580 (Fax)

May 2010

I thank the National Science Foundation Graduate Research Fellowship program for financial assistance, the National Center for Supercomputing Applications for computer use and Natalie Ko and Nat Keohane for data. I also thank R. J. Briggs, Russ Cooper, Dean Corbae, Don Fullerton, Arik Levinson, Rob Stavins, two anonymous referees, and seminar participants at Texas, Maryland, Stanford, Princeton, Brown, Washington, Northwestern, Harvard, Georgetown, the University of California Energy Institute, the Econometric Society meetings in Minneapolis, and Camp Resources XIII in Wilmington, NC, for helpful comments.

Abstract

Environmental regulations that grandfather existing plants by not holding them to the same strict standards as new plants may have the unintended consequence of retarding new investment. If new plants are cleaner, then this effect may increase pollution in the short run. I develop a dynamic model of a facility's decisions over scrapping and abatement, which depend on capital depreciation, profitability shocks, and environmental policy. Using data from fossil fuel fired boilers at electric power plants, I estimate the structural parameters of the model and assess the impact of grandfathering in the Clean Air Act on sulfur dioxide emissions. Counterfactual policy simulations show that an increase in the stringency of performance standards would have led to a decrease in investment in new boilers. However, this does not lead to increased emissions, since there is less investment in dirtier coal boilers as compared to relatively cleaner oil or natural gas boilers.

JEL Codes: Q58, L94

Keywords: Clean Air Act, Sulfur Dioxide, Method of Simulated Moments, Vintage-Differentiated Regulation, Electric Power Plants

Regulations often contain grandfathering provisions, where facilities already built or workers already employed at the time of passage are not subject to the new standard. While the reasoning for such provisions may relate to fairness, such as a wish not to "change the rules in the middle of the game," they often come with unintended consequences. By giving different incentives to grandfathered agents and non-grandfathered agents, the regulations can lead to unexpected outcomes. The federal Clean Air Act (CAA) and its New Source Performance Standards (NSPS) for major sources of air pollution are examples. By mandating that any new pollution sources (including power plants) meet strict pollution control standards, the rule may keep older facilities from scrapping and being replaced by newer ones. If older facilities are dirtier than newer ones, then this may create a "perverse" effect that increases pollution.

The purpose of this paper is to determine how grandfathering provisions in environmental policy affect both investment in and pollution from fossil fuel fired electric power plants. I develop a dynamic model of each facility's decisions about investment in new capital and choice of abatement techniques. These decisions are affected by the relative profitability of new capital, the costs of upgrading, and environmental regulations. Newer capital pollutes less, and hence stricter environmental policy without grandfathering provides an extra incentive to upgrade. Yet stricter environmental policy with grandfathering may provide a disincentive to upgrade. Using data from 1985-1995 on U.S. electric power plants, I estimate the parameters of the model. Finally, I simulate the estimated model to determine how grandfathering in the CAA impacts emissions and investment.

An early examination of a grandfathered environmental policy is [1], who looks not at stationary air pollution sources but at automobiles. He finds that stricter vehicle emissions standards, which apply only to new cars and hence effectively grandfather old cars, lead to a perverse short term increase in emissions. [2] find that environmental regulations increase the age of capital but not the level of emissions for electric utilities, while [3] finds no significant difference in capital vintage between states with and without grandfather provisions for manufacturing plants in commercial printing and paint manufacturing. [4] and [5] also find perverse effects of grandfathering in the electric power industry and in manufacturing plants in New York state, respectively. Finally, [6] find that grandfathering in the CAA decreases the capital expenditures of coal-fired power plants but has no effect on their operating costs, fuel efficiency, or emissions.

In this paper's model, firms face a discrete choice of whether or not to scrap and upgrade their capital. This model contains two important extensions to discrete choice scrapping models to accommodate this industry and this policy. First, since the regulations are grandfathered, the model contains an additional state variable representing a facility's grandfathered status. Investment affects this status (if you scrap your plant, you lose its grandfathered status), so the policy impacts dynamic investment decisions. Second, instead of just considering a single binary decision (scrap or don't scrap), the model includes decisions over pollution abatement and plant type. These are important avenues for plants to respond to environmental policy, and without including them as options, plants are constrained in how they can react to policy changes.

This paper also extends the literature that studies the effects of grandfathering provisions in the CAA, including NSPS and New Source Review (NSR), by developing the first model that is structural and dynamic. A structural model allows for simulation of countless counterfactual policy scenarios. A dynamic model makes explicit the dynamic programming problem that utilities are solving, and it captures the effect that the cross-sectional distribution of plant vintages and grandfathering statuses has on the impact of the policy. Although this paper uses the model to retrospectively analyze the CAA's effect on the electricity industry through the mid-1990s, the model can be adapted to study and make forecasts about other industries and other policy environments in which grandfathering plays a role.

Grandfathering, or vintage-differentiated regulation, appears frequently outside of the CAA, both in environmental and non-environmental laws. Corporate average fuel economy (CAFE) standards and emissions rate standards apply only to new cars, so old cars are effectively grandfathered. The Clean Water Act and the Safe Drinking Water Act both set differential standards for water treatment plants based on when they went into operation. Furthermore, grandfathering is by no means absent from current policy debate.[1] Fire sprinklers are required in new buildings, but existing buildings are often not required to have them unless they are renovated. Zoning ordinances generally do not apply to businesses or homes built before the ordinance went into effect. Given the prevalence of grandfathering, it is important to study its effects, including any potentially perverse or counterproductive effects.

I find a significant effect of grandfathering in environmental regulations on both emissions and investment. Using the model to simulate the CAA and counterfactuals, I find that if grandfathering provisions were eliminated in 1985, emissions from power plants would be 60% less than the baseline level in 1995. This is because a large majority of plants in the cross section are grandfathered, and eliminating grandfathering amounts to a large policy shift. If the stringency of the grandfathered CAA standards was weakened in 1985, investment in new boilers would rise and yet emissions would also rise. The rise in new investment occurs for the same reason that strengthening a grandfathered standard can reduce new investment. When the grandfathered regulations are weakened, grandfathered plants lose some value in their grandfathered status, and hence their disincentive to invest falls. However, emissions do not fall, because the increased new investment occurs in relatively dirty coal boilers as opposed to relatively clean oil or natural gas boilers. Thus, the structural model that includes choice of boiler type helps explain results from previous studies, including [2] and [6], finding that grandfathered regulations inhibit new plant investment but have no perverse effect on emissions.

The next section below presents the model. In section two I describe the data. Section three presents the estimation strategy and results, section four the results from the counterfactual simulations, and section five concludes.

I. Model

I develop a model to estimate the impact of grandfathering in the Clean Air Act of 1970 on the electric power generating industry.[2] The CAA is an appropriate policy to consider because of its explicit grandfathering of existing sources. Section 111 of the law gave EPA the power to set binding emissions standards on all new sources of emissions - the New Source Performance Standards. Regulation of existing plants was left up to the states and is likely to be less strict. The plants were grandfathered for reasons of efficiency (it is costlier to retrofit existing plants than new ones), equity (it is unfair to "change the rules of the game mid-stream" by regulating existing plants), and politics (potential facilities have less clout than existing ones).[3]

I consider the behavior of a profit-maximizing one-plant firm. This model could be generalized to multi-plant firms, but the assumption that each individual plant's maximization decision is independent leads to identical results. The basis of this model is the discrete choice scrapping model, e.g. [7], extended in two important ways. First, in addition to the choice of scrapping, plants can choose abatement investment to reduce emissions and can choose among different types of capital. Second, plants can be grandfathered from environmental policy.

Five nonrandom and two random state variables describe the plant at each period. First, the plant's age is v. Second, a plant can be of two types: coal (b = 1) or non-coal (b = 0).[4] Third, a coal plant can have a flue gas desulfurization unit, or "scrubber," to reduce its emissions of sulfur dioxide (SO_2). If so, then $s = 1$, otherwise $s = 0$. Fourth, a plant can be grandfathered from environmental policy. Let $u = 0$ if the plant is grandfathered and exempt from the policy, and $u = 1$ if not. Fifth, a plant faces a price differential, p, between the prices it faces for using low-sulfur coal and high-sulfur coal.[5] Finally, two random state variables, A and B, represent a productivity shock and a shock to emissions, respectively. The state variables are summarized in the first column of Table 1.

Each plant in each period faces decisions over four variables. First, it makes a discrete decision whether or not to update its production technology, that is, whether or not to scrap and replace with a new plant of age 1.[6] If updating, it can be replaced with a coal or a non-coal plant. Let the decision variable $z = 0$ if no replacement, = 1 if replaced with a coal plant, and = 2 if replaced with a non-coal plant. The second decision faced by plants is whether to add a scrubber. This decision is possible only for coal plants. Let $x = 1$ if the plant chooses to add a scrubber and $x = 0$ otherwise. Third, also available for coal plants only, is the opportunity to use low-sulfur "clean" coal as a means of reducing SO_2 emissions. Let $c = 1$ if the plant chooses to use clean coal and 0 otherwise. Fourth, the plant faces a continuous decision over the intensity of use. This is measured by the amount of fuel input, m, used by the plant in a period. The second column of Table 1 summarizes the four choice variables.

A plant's single-period profit function is given by

$$Af(v,m,b) - F_c \cdot 1(z{=}1) - F_{nc} \cdot 1(z{=}2) - xG - pcm - \tau ue/m.$$

The first term, $Af(v,m,b)$, represents the reduced-form revenue of a plant of age v, operating at intensity m, and of type b. Here A is a multiplicative productivity shock. By reduced-form revenue I mean that the utilities are implicitly optimizing given supply and demand curves and whatever regulations they face, including rate of return regulations on electricity prices. Given the conditions under which these utilities operate, some function exists that they are trying to optimize, and it varies by capital age, type, and intensity of use. The next two terms capture fixed adjustment costs, F_c and F_{nc}, that are faced only if a plant is replaced with a coal plant ($z = 1$) or a non-coal plant ($z = 2$), respectively. G is the fixed cost of adding a scrubber, paid only when a scrubber is added ($x = 1$). The per-unit cost differential between clean and dirty coal is p, so that when clean coal is used ($c = 1$), p is paid per unit of coal m.[7] This price differential p is allowed to vary by plant and over time.

The last term in the expression for profit, $-\tau ue/m$, represents the implicit costs of environmental policy. A plant's emissions is given by e. Environmental policy is modeled by a virtual tax rate τ levied on emissions intensity e/m. The tax payment $\tau e/m$ is paid only if the plant is not grandfathered, that is, $u = 1$. If the plant is grandfathered, then $u = 0$ and no environmental tax is paid.

The NSPS are performance standards, not an emissions tax, though here they are modeled as a tax. These standards create an implicit price on emissions, and utilities respond to the NSPS the same as they would respond to a tax.[8] Importantly, the tax is levied not on the quantity of emissions e but on the emissions intensity, e/m. The NSPS have always been targeted towards emissions intensity rather than absolute emissions. For example, the NSPS for new coal plants set in the 1970 CAA was 1.2 lbs of SO_2 per MMbtu of heat input. By modeling the NSPS as a tax on emissions intensity, plants cannot limit the impact of this tax by reducing operating intensity only while maintaining the same ratio of emissions to fuel input. Thus, it more closely matches the regulatory environment these firms faced; no evidence exists suggesting that plants met their NSPS through reduced operating intensity.

Different specifications of the policy have different implications for firm behavior. Many observers have claimed that regulations rewritten in 1978 effectively required the installation of scrubbers on all new coal plants (although the data show that not all plants built immediately after 1978 did so). In a sensitivity analysis, I consider policy modeled as a tax on the level of emissions rather than emissions intensity.

Emissions, like revenue, are a function of plant age, type, and intensity of use, but also of abatement activity: $e = Bg(v,c,s,m,b)$, where B is a random shock to emissions. Like a total factor productivity shock, this shock could arise from an innovation in abatement technology. It is expected that older plants emit more, plants using scrubbers or clean coal emit less, and a higher operating intensity leads to higher emissions; all of these properties are found in the data.

The final column of Table 1 describes the transition process between states. If a plant of age v updates, its age next period (v') becomes 1. If it does not update its next-period age is $v + 1$. The plant type next period b' is determined by the replacement decision. The indicator for the presence of a scrubber next period, s', equals one if either a scrubber was present this period and the plant was not replaced or a scrubber was built this period. If a plant updates, then it loses its grandfathering status so $u' = 1$. The last three state variables, p, A, and B, are determined exogenously by an as-yet unspecified Markov process, so that the expectation of next periods' values of these variables is taken as a function of this process.

Plants maximize discounted expected lifetime profits over an infinite horizon $E_0 \sum_{t=0}^{\infty} \beta^t y_t$ where E_0 is the expectation operator at $t = 0$, β is the discount factor, and y_t is profit in period t. The plant's choice can thus be written as a dynamic programming problem:

$$V(v,s,u,b,p,A,B) = \max_{\substack{m \in [0,\overline{m}] \\ z \in \{0,1,2\} \\ x \in \{0,1\} \\ c \in \{0,1\}}} \left\{ \begin{matrix} Af(v,m,b) - F_c \cdot 1(z=1) - F_{nc} \cdot 1(z=2) - u\tau Bg(v,c,s,m,b)/m \\ -xG - pcm + E_{A',B',p'|A,B,p} \beta V(v',s',u',b',p',A',B') \end{matrix} \right\},$$

where the state variables are determined according to the transition processes described in Table 1. Even when parameterizing the revenue function, the emissions function, and the process governing the evolution of the shocks and p, this model cannot be solved analytically. The discreteness of three of the four choice variables makes it impossible to create a series of first-order conditions. Thus, the model is solved numerically through value function iteration (VFI). The Appendix details this method.

While this model does not explicitly incorporate technological growth, it can be derived from a model with an exogenous growth rate that is factored out to make the problem stationary.[9] Suppose that the productivity of a new plant increases at the rate κ each year. Older capital is less productive both absolutely, because of real depreciation at rate, say, δ_r, and relatively, because of the improved technology of newer capital. Thus "capital depreciation" in this model encompasses both physical depreciation (δ_r) and obsolescence (κ). This model cannot separately identify δ_r from κ, but it includes both. Since I am concerned with analyzing policy by simulating counterfactuals, this does not present a problem so long as these two forms of depreciation remain constant under the various counterfactuals. Thus, the model does not accommodate *endogenous* technological growth, which may be induced by changes in environmental policy. A counterfactual strengthened environmental policy may induce additional growth in abatement technology, and thus the counterfactual simulation here will overstate emissions from that counterfactual. So, long-run predictions of the model should be viewed with more caution than short- or medium-run predictions, where technological change cannot yet respond to policy. In the simulations I analyze the first ten years after the policy change.

While federal emissions regulations have been grandfathered since the 1970 CAA, Title IV of the 1990 Clean Air Act Amendments (CAAA) established an emissions trading program for SO_2 emissions among electric utilities beginning in 1995.[10] Because of this regulatory change, I do not estimate the model on data after 1995. This policy was not grandfathered; plants were required to participate regardless of their vintage. Therefore, the analysis here should be viewed as a retrospective analysis of NSPS for SO_2, since those standards exist alongside the permit trading market. The extent to which plants respond in anticipation of a forthcoming policy (for example, by installing a scrubber in 1992 so that their emissions can be reduced when the policy takes effect in 1995) can be captured in this model by incorporating dynamic policy.[11] Another potential concern is the restructuring of electricity markets, though I suspect that this will be unimportant during my sample period (1985-1995), since restructuring did not begin until the late 1990s.[12]

A different grandfathered policy created in the 1977 CAAA, New Source Review (NSR), applies to both new plants and existing plants making major modifications. Thus, a unit could be built before 1970 and hence be grandfathered from NSPS, but have made major modifications that triggered NSR and hence be subject to federal environmental regulations. As identified by [6], this additional aspect of NSR creates a second distortion from grandfathering: in addition to keeping older plants from scrapping, it keeps them from making modifications that might trigger

NSR, although these modifications may increase productivity and reduce emissions. The second distortion is unlikely to actually apply to any units in my sample period (1985-1995), since it has been widely documented that there had been no vigorous enforcement of this aspect of NSR prior to the late 1990s.[13]

II. Data

The econometric analysis is done at the boiler level as opposed to the plant level. Each fossil fuel fired power plant can have multiple boilers that could be of different vintages, have different grandfathered statuses, and use different abatement strategies. The data come from three sources. The Energy Information Administration's Form EIA-767 contains information on each boiler, including age, intensity of use, and abatement choices, and is used to calculate the emissions function for coal boilers and the moments for the structural estimation. Form EIA-423 contains plant-level fuel prices for both clean and dirty coal as well as oil and natural gas. The EPA's Emissions and Generation Resource Integrated Database (eGRID) provides emissions data used in estimating the emissions function for non-coal boilers.

Emissions data are not available for the years studied (1985-1995). To estimate the emissions function, I use two different strategies, one for coal boilers and one for non-coal boilers. For coal boilers, information from the Form EIA-767 is used to calculate each boiler's annual emissions from its amount of coal used, average sulfur content of coal, and abatement technology. An engineering equation is used to calculate emissions from these data on a mass-balance basis, as in [8].[14] Given these emissions estimates (in tons of SO_2), the emissions function $g(v,m,s,c,b)$ for coal boilers is estimated using OLS. These emissions calculations are only available for coal boilers. For non-coal boilers, I use measured emissions data from eGRID.[15] While measured emissions data ought to be preferable to those calculated from an engineering equation, the measured eGRID data are only available from 1996-2000, after the sample period.[16]

Data on boiler age, total fuel input, use of clean coal and scrubbers, and grandfathering status come from the EIA-767 data files.[17] These data cover steam-electric plants (including fossil fuel fired plants) with a generator nameplate rating of 10 or more megawatts. I use data from 1985-1995, which contain approximately 1800 boilers per year.

Boiler age is calculated from the inservice date reported, or the year in which the boiler went into operation. The grandfathered status of a boiler, that is, whether it is subject to NSPS, is reported in the data set.[18] In 1985, 87% of boilers are grandfathered from NSPS. Figure 1 presents a histogram of the vintage distribution of all operating boilers in the 1985 EIA-767 data file, categorized according to grandfathered status. While virtually all boilers built before 1970 are grandfathered, and all built after 1978 are not, during the intermediate period variation exists in grandfathering status even for boilers of the same age.

The EIA-767 data also provide information on abatement activity: scrubbers and clean coal. In 1985, 15% of coal boilers used a scrubber. The data include the average sulfur content of the coal used, which can be converted to pounds of sulfur per MMbtu of heat input. This value is continuous, as shown in Figure 2, a histogram of the distribution of average sulfur content of coal for 1985 coal boilers. For simplicity, I model the choice of coal type as a binary variable; it is either clean (low-sulfur) or dirty (high-sulfur). Following [9], I choose the cut-off point between the two coal types to be 1.2 lb/MMbtu, the value of the performance standard for new boilers set in the 1970 CAA. Based on this definition, 34% of coal boilers use clean coal in 1985.

This differential is clearly due to heterogeneous price differentials between clean and dirty coal. The EIA-423 data contain plant-level information on total quantity and average price of annual coal consumption. [19] The data are given for the price actually paid for the coal actually received by the plants. A single plant can use coal from multiple sources, some of which may be high-sulfur and some of which may be low-sulfur. For plants that use both types of coal, I am able to construct an average price paid for each of the two types. Many plants, though, only use one type of coal. For these plants, I use the state average of coal prices for the other type of coal. For some states, no plants use a particular type of coal, and I use the average price for the census region of a particular type of coal for plants in these states. For those plants with multiple boilers, I assign each boiler the same plant-level value of coal prices.

Coal price data are available for coal plants only. Since the model allows for non-coal boilers to scrap and be replaced by coal boilers, I also want coal prices available to these boilers. Thus for non-coal boilers I assign coal prices based on the state or regional average coal prices.

The last item obtained from the EIA-767 data set is the annual amount of fuel input used by a boiler. For each fuel, I convert the fuel amounts into the amount of heat energy (in MMbtus). While 58% of all boilers are coal in 1985, they represent over 90% of the total amount of fuel input. Most of the remaining 10% is from gas boilers.

In addition to the differential fuel use by type of boiler, the age of the boiler also makes a large difference in the operating intensity. Boilers built after 1970 are used much more intensively. The grandfathered boilers represent a large fraction of the industry by number (85%) but only a small fraction by fuel use (7%). Boiler efficiency is declining in age, averaging 86.4% for the youngest decile of boilers and 84.0% for the oldest decile.

The moments used to identify the parameters in the structural portion of the estimation are presented in Table 2. I choose two sets of moments. First, I use a set of moments from each year in the sample period. Table 2 includes moments from the first year, 1985, although they will not be used in the estimation, since the 1985 data are used to set the initial distribution of boilers. Also, I omit 1995, the last year for which I have data, since one of the moments requires data from the following year to construct it. For each year in 1985-1994, I calculate the fraction of boilers grandfathered, the fraction that uses coal, and of the coal boilers, the fraction that has a scrubber and the fraction that uses clean coal. These annual data moments are meant to identify changes in the industry over the different years and match how boiler characteristics change with time. Second, I aggregate the boilers from all ten years of data and divide them into five categories based on their age and grandfathered status: less than or equal to 15 years old, between 16 and 25 years old and grandfathered, between 16 and 25 years old and not grandfathered, between 26 and 35 years old, and greater than 35 years old. The age group between 16 and 25 years old is divided into grandfathered and non-grandfathered boilers since it is those ages of boilers for which there is variation in the grandfathered status. For each category, I evaluate the fraction of boilers scrapping, the fraction of boilers that uses coal, of the coal boilers, the fraction with scrubbers and the fraction using clean coal, and the fraction of boilers in each category. These categorical data moments are used to identify the impact of both age and policy on boiler operating decisions.

The top half of Table 2 displays the annual moments. The fraction of boilers grandfathered declines slightly over the period from 87% to 84%. Of coal boilers, the fraction using scrubbers increases from 14% to 18%. The percentage using clean coal goes from 34% in

1985 to 41% in 1994, but not monotonically; this fraction peaks at 46% in 1988. The fraction of all boilers that are coal boilers increases slightly from 58% to 62%.

The second panel of Table 2 displays the categorical moments. While the annual moments do not differ significantly from year to year, the categorical moments do strikingly differ by category. The first row shows that the fraction of generators scrapping is very low. The percentages in the first four categories are all less than three percent, and they show that a very small number of boilers are scrapping. For example, the value of 0.5265% for boilers younger than 15 years old comes from 18 boilers scrapping out of 3,419 in that age category over the ten years. The patterns in the data are conformable to intuition. One would expect that grandfathered boilers are less likely to scrap than non-grandfathered boilers, and that is just what is seen comparing the second and third categories. As boilers get older, they are more likely to scrap; 7% of boilers older than 35 years old scrap.

The moments related to abatement activity are consistent with boilers responding to grandfathered policy. Grandfathered units are less likely to have a scrubber or to use clean coal. The rate at which boilers adopt either of these abatement measures decreases with boiler age. The youngest boilers are more likely to be coal fired, but there is no difference in the fraction of coal boilers between the other age categories. About 35% of the boilers are more than 35 years old, while about 16% are younger than 16 years old.

III. Estimation Strategy and Results

The model is estimated in three steps. First, the emissions function $g(v,m,s,c,b)$ is estimated with OLS, after imposing a functional form. This equation is an emissions production function; given the plant's age, abatement equipment, type, and choice of operating intensity, a level of emissions will be produced. Second, I estimate the revenue function $f(v,m,b)$ using panel data estimation allowing for boiler-specific effects and autocorrelated error terms. Finally, since it is impossible to estimate all of the parameters in this way, I estimate the rest of the parameters using the method of simulated moments (MSM).[20] While all of the parameters could be estimated using MSM, estimating a subset of parameters using other methods when possible is preferable because it saves computational time and the identification of those parameters is more straightforward.

Because the emissions function, g, is a function of plant type, b, I estimate a separate emissions function for coal and non-coal plants. The estimating equation for coal boilers is

$$ln(emis) = \beta_0 + \beta_1\, ln(age) + \beta_2\, ln(m) + \beta_3\, lsc + \beta_4\, scrubber + \beta_5 lsc*scrubber + \varepsilon.$$

The age of the boiler, the amount of fuel input (m), whether the boiler uses low sulfur coal (lsc) and whether the boiler has a scrubber are all included, as is the interaction term between the last two abatement methods. The estimating equation for non-coal boilers is identical, save that the *lsc* and *scrubber* terms (and their interaction) are not included. I do not estimate the emissions shock B because I do not have actual emissions data and thus do not have to accommodate variation in the data.

Results are presented in Table 3. The first four columns are the results for coal boilers from the calculated emissions data from 1985-1995. Unreported state and year dummies are included in all columns of Table 3 to account for unobserved heterogeneity in the regulatory environment at the state level, as well as trends in emissions picked up by the year effects. Older plants emit more, once controlling for total heat input and abatement activity. The elasticity of emissions with respect to age is about 0.13. Note that this is measuring both physical depreciation and technological obsolescence; older boilers are both absolutely and relatively dirtier. The use of scrubbers and the use of low-sulfur coal significantly decrease emissions, as expected, and the interaction term is positive (since a scrubber has less to "scrub" when low-sulfur coal is already being used). The elasticity of emissions with respect to heat input is almost exactly one. This suggests that emissions are linear in heat input, which is used in identifying the revenue function off the first order condition.[21]

The next column of Table 3 uses the log of the emissions rate, in tons of SO_2 per MMbtu heat input, as the dependent variable. This estimating equation is

$$ln(emis/m) = \beta_0 + \beta_1\, ln(age) + \beta_2 scrubber + \beta_3\, lsc + \beta_4\, lsc*scrubber + \varepsilon.$$

The emissions rate is also higher for older boilers, with an elasticity of about 0.14. Scrubbers and low-sulfur coal have the expected effects.

Columns 3-5 consider alternate specifications for coal boilers. In each of these columns the dependent variable is the log of the emissions rate. Column 3 includes an indicator variable equal to one if the boiler's county is not in attainment of the Clean Air Act's air quality standards for SO_2. It also includes a measure of relative prices of three fossil fuel inputs, averaged at the state-year level from EIA-423 data. Column 3 also removes all observations from states that had begun electricity restructuring by 2000, to eliminate any potential bias from utilities' anticipatory behavior of restructuring. Boilers in nonattainment counties have a significantly lower emissions rate, and fuel prices do not enter significantly. The coefficient on the log of boiler age is 0.17, only slightly higher than in column 2.

Column 4 includes three additional variables related to boiler technology or intensity. These are the capacity of the boiler's plant (in megawatts), the heat rate of the boiler's plant (in BTU/kilowatt-hours), and the number of hours the boiler is in operation over the year. Plant capacity and hours of operation enter negatively, while heat rate is insignificant. The coefficient on the log of boiler age is still significantly positive, but falls to 0.05. Lastly for coal boilers, column 5 performs the same regression as column 2 but uses the actual emissions data from eGrid (available only from 1996-2000) rather than the estimated emissions data from EIA-767 (from 1985-1995). The coefficient on the log of boiler age is still significant but is larger than in column 2.

The regression results for non-coal boilers are presented in columns 6-9 of Table 3. Column 6 regresses the log of total emissions on the log of age and the log of total heat input, as well as unreported state and year fixed effects, and column 7's dependent variable is the log of the emissions rate. Column 8 includes county-level attainment status, relative fuel prices, and omits states that began restructuring by 2000; column 9 includes capacity, heat rate, and hours. The non-coal regressions do not tell a consistent story. Only in column 8 is there a positive and significant coefficient on age, as expected; column 7 even finds a negative (though insignificant) coefficient.

This unusual result is not worrisome for two reasons. First, the contribution of non-coal boilers to SO_2 emissions is small; coal plants are responsible for 97% of total SO_2 emissions from this sector. If the non-coal regression results in this stage are off, it will not substantially affect the simulations. Second, the absence of a positive elasticity between age and emissions for non-coal plants is possible. Technological innovation in abatement of SO_2 emissions from power plants has focused mainly on coal plants because of their dominance in this pollutant. Since the age elasticities in Table 3 include the effect of obsolescence from technological growth, the lack of a positive age effect may signal this lack of technological growth in this area. I use the point estimates from columns 2 and 7 in the MSM estimation and simulations.

In the second stage of the estimation, I exploit the fact that two of the choice variables are chosen in a static, not dynamic, problem. Both m, the operating intensity, and c, the sulfur content of coal input, affect the single-period profit function but no future state variables and thus not the next period value function.[22] The optimal decision over these two variables can be made by just looking at the single-period profit function. In the case of m, a continuous choice variable, a first-order condition arises:

$$Af_m(v,m,b) - \frac{d}{dm}[u\tau Bg(v,c,s,m,b)/m] - pc = 0\,.$$

The first term in the equation represents the marginal benefit to a plant of an additional unit of fuel input from increased revenue. The middle term represents the costs from the environmental policy as measured by a marginal increase in the virtual emissions tax paid, paid only by non-grandfathered plants ($u = 1$). The last term is the marginal cost of using clean rather than dirty coal.

The middle term of the first-order condition disappears when the emissions function g is linear in fuel input m. In the emissions estimates above, this is just what is found. Thus, this moment condition is simplified since it is now independent of the emissions function. It depends on the parameterization of f, defined as follows:

$$f(v,m,b = coal) = d_{0c}m + d_{1c}vm - ½\, m^2$$

$$f(v,m,b = non\text{-}coal) = d_{0nc}m + d_{1nc}vm - ½\, m^2.$$

This parameterization captures an important dimension of the behavior of these boilers: older plants are used less intensively. Under this parameterization, this behavior of plants can be captured when d_{1c} and d_{1nc} are negative. The profit function must somewhere be normalized for identification; thus I fix and do not estimate the coefficient on m^2.

Given this parameterization, the first order condition for fuel use for non-coal boilers is $d_{0nc} + d_{1nc}v - m = 0$. The first order condition for coal boilers is different, since they must also optimize over the binary choice of choosing clean or dirty coal, given the price differential p. That condition is $d_{0c} + d_{1c}v - m - pc = 0$. The emissions function for coal plants has the additional variable c for low sulfur coal use. Also, the per-unit price differential appears in the condition. A coal boiler optimizes over whether to use clean or dirty coal (c = 1 or 0), but after doing so the boiler chooses a level of fuel input which satisfies the first order condition above, for either value of c.

The functional form chosen thus allows each revenue function to be estimated via a least-squares regression, where the independent variables are boiler age and a constant term, and the dependent variable is fuel input m. For coal boilers, the regression also includes the clean-dirty coal price differential times an indicator for the use of clean coal (pc). Because of the normalization, the revenue function and indeed the entire profit function are only identified up to scale. Thus, I estimate a scale coefficient on the pc term.

The reduced-form revenue function abstracts from supply and demand curves and from the strategic behavior of firms operating under regulators. Because markets as well as regulations differ by time and location, I include in the revenue function estimation state- and year-fixed effects.

These results are presented in Table 4. Columns 1–4 present the results for non-coal boilers, and columns 5–8 present the results for coal boilers. Columns 1 and 5 are simply OLS regressions with standard errors clustered at the boiler level; the significant negative coefficient on age indicates that older boilers are used less intensively. For coal boilers, the coefficient on pc is negative, as expected, though only significantly so in columns 6 and 7.

Columns 2 and 6 use a random-effects specification, to pick up an unobserved boiler-specific effect. These specifications result in a coefficient on age that is still significantly negative but has a smaller magnitude, by about a factor of one-half for non-coal boilers and a factor of two-thirds for coal boilers. Columns 3 and 7 keep the random-effects specification and also add state-year relative fuel prices as regressors. Given that many boilers and plants belong to a utility's fleet and are dispatched on a least-cost basis, relative fuel prices may matter. The coefficient on age does not change after inclusion of the price variables. Finally, in columns 4 and 8 the error is allowed to be autocorrelated according to an AR(1) process; again the coefficient of interest is largely unchanged. The preferred specification for the parameter values comes from columns 4 and 8, and those estimated parameters are used in the simulations.

Finally, in the third stage I use MSM to estimate the remaining parameters of the model. This method is described in more detail in the Appendix. These parameters include the fixed costs for building a coal and non-coal boiler, F_c and F_{nc}, and the fixed cost of building a scrubber G. The implicit tax on emissions intensity τ is also estimated in this stage. Ideally, this could be calibrated based on the policy. Because the complex CAA is modeled simply as an emissions tax, where the value of the tax is the shadow price on pollution intensity that the policy creates, this value is unknown. Hence, it falls into the parameter set to be estimated. The productivity shock A is idiosyncratic.[23] It is assumed to be multiplicative and distributed log-normally with median one. That fixes its first parameter, μ, to zero, while σ_A^2 is estimated. The shock is persistent and evolves according to a Markov process. For simplicity, I allow the transition matrix to be defined by one parameter, P. A boiler has probability P of having the same productivity shock in the next period. With probability $1 - P$, the next-period productivity shock is randomly chosen from the log-normal distribution.

The annual discount rate β is set at 0.95.[24] This leaves six parameters to be estimated with MSM: $[F_c, F_{nc}, G, \tau, \sigma_A^2, P]$. The data moments that are used in MSM are described in the following section and listed in Table 2.

The MSM estimation differs from estimation techniques used in previous literature related to grandfathering and the CAA in that the technique here is the first that is both structural and dynamic. Previous papers have identified a perverse effect of grandfathering in a reduced-form setting. Here, I develop and estimate a structural model and use it to quantify the effects of policy changes in the presence of grandfathering. Furthermore, the model is explicitly dynamic, while many of the reduced form models are static, using cross-sectional data. The effects of grandfathered regulations are inherently dynamic; firms or individuals delay turnover to take advantage of the grandfathered exemptions in the future. Though this dynamic model is implicit in those reduced form models, it is made explicit here.[25] As a result, the mechanisms by which policy affects firm behavior can be better understood.

As with any estimation but in particular with structural estimation, one must pay careful attention to identification of the model's parameters. In the MSM estimation stage, I identify the implicit tax on emissions using variation in the behavior of grandfathered versus non-grandfathered facilities. Grandfathering is based on age, so it is difficult to separate an age effect from a grandfathering effect. I do so by exploiting the fact that of the facilities aged 15-25 years, some are grandfathered and some are not. It is unfortunately impossible to determine the source of this variation from the data. The age is based on the date the plant went online, but the grandfathering status for some plants may have been partly based on the date of construction. For some facilities, the grandfathering status could be determined by the utilities bargaining with regulators to adjust the official online date.

Clearly, grandfathering status was determined by someone and not randomly assigned. The question is how or if this determination creates a bias. Suppose that of the plants within that age group, those with the highest expected abatement costs were more likely to be grandfathered, say as a result of lobbying. The higher use of abatement among non-grandfathered plants is thus partly due to the unobservable difference in abatement costs, though the estimation strategy attributes it to the policy. Thus the estimated implicit tax created by the policy may be biased upwards. Alternatively, if the selection into grandfathering based on abatement costs went the other way, then there is a downward bias on the estimate. It is not possible to measure the direction or magnitude of this potential bias.

The results from the MSM estimation are presented in Table 5. Each point estimate is presented with its estimated standard error, and all parameters have small standard errors save σ_A^2. The fixed cost parameters F have no units, so an estimated value of 1800 does not represent 1800 dollars or 1800 million dollars. Rather, the magnitude is normalized to the boiler's profit function $y = Af(v,m,b) - F_c \cdot 1(z=1) - F_{nc} \cdot 1(z=2) - \tau ue/m - xG - pcm$. The relative magnitudes of the three parameters give information about construction costs. The cost of a new coal boiler is 20 times as large as that of a new non-coal boiler; coal boilers and plants tend to be of much larger capacity, and this is represented both in the profit function of the boiler and in the fixed cost of building one. The cost of adding a scrubber to a coal boiler is about one-thirtieth the cost of the boiler itself. These relative costs corroborate engineering estimates of production costs.

Like the fixed costs, the policy parameter τ can be interpreted in relation to the profit function. The emissions function g gives the emissions intensity in tons of SO_2 per MMbtu fuel input as a function of age and abatement technology. The implicit tax τ is levied on emissions intensity. The amount that is paid by a boiler in the virtual tax (if it is non-grandfathered) is the product of the estimated value of τ and its emissions intensity. For example, a coal boiler aged 20, with no scrubber and not using low-sulfur coal, has an estimated emissions intensity of 0.0016 tons SO_2 per MMbtu fuel input. At the estimated tax value of about 3000, this equates to a tax payment of 4.8. Though this has no units, it can again be compared to the other parameters. Consider the cost of a scrubber, 61.41. This is about 13 times the annual payment in implicit emissions tax for that particular boiler. If that same boiler had a scrubber, its emissions intensity would be 0.000442, and its tax payment would be reduced to 1.3. Thus, utilities are balancing long-run projections of emissions tax payments with scrubber costs to decide when to add a scrubber.

The fact that τ is significantly different from zero amounts to a rejection of the null hypothesis that grandfathering has no effect on the behavior of plants. Since the implicit tax is paid only by non-grandfathered boilers, the significantly positive tax means that these boilers respond to an incentive to reduce their emissions that the grandfathered boilers do not.

The parameter $\sigma_A{}^2$ gives information about the variance of the random productivity shock. The estimated value of the parameter, about 0.1104, corresponds to a variance of 0.4588.[26] Finally, P represents the persistence in the idiosyncratic component of the productivity shock. If $P = 0$, then no persistence exists, and if $P = 1$ then the idiosyncratic component is constant. Here, P is 0.87. A generator has an 87% chance of staying in its current productivity state and a 13% chance of taking a random draw from the distribution of shocks; the idiosyncratic element of revenue is quite persistent.

The Appendix compares the simulated moments to those from the data; in general they match quite closely.

IV. Counterfactual Simulations

Finally, I use the estimated parameters to simulate the effects of policy changes. Specifically, I consider the impact of grandfathering in the CAA by comparing a baseline simulation using the estimated parameters to three counterfactual simulations. In all counterfactuals I present the simulation results for ten periods after the policy change, though the simulations come from an infinite horizon maximization model. In the first counterfactual, I keep all parameters the same, except that I eliminate any grandfathering provisions starting in the first year. Next, I set the implicit tax rate on pollution equal to 90% of its estimated value to simulate what would have happened had the stringency of the grandfathered CAA regulations been reduced by 10% in 1985. Finally, the third counterfactual increases the implicit tax rate on pollution by 10%, to represent the CAA regulations increasing in strength in 1985. The counterfactual that eliminates grandfathering is chosen to quantify the effects of grandfathering itself in the policy, rather than the effects of the stringency of the policy. The two other counterfactuals, which keep grandfathering but alter the stringency of the policy for non-grandfathered units, are chosen to investigate how policy changes that maintain grandfathering will affect outcomes. This is an exercise similar to that in [1], which looks at the effect of changes in automobile emissions standards that are grandfathered.[27]

For the eleven periods presented of the three counterfactual simulations, Figure 3 presents the level of emissions and Figure 4 shows the number of boilers that adjust in each period. The line "No Grand." represents the simulation that eliminates grandfathering. The line "Tax Cut" represents the simulation with the tax rate reduced. The line "Tax Hike" represents the simulation with an implicit tax rate increased by 10%.[28] In Figure 3, the emissions levels are presented as proportional deviations from the baseline simulation. For example, in 1995 emissions under the "Tax Hike" counterfactual are 90% of emissions under the baseline scenario. In Figure 4, the number of boilers adjusting is presented as the absolute deviation from the numbers in the baseline simulation. For example, in 1985, 50 more boilers scrapped under the "Tax Cut" counterfactual than under the baseline scenario.

Consider first the "No Grand." simulation, which shows the largest difference from the baseline for both emissions and scrapping. In Figure 3, emissions drop sharply. Since older boilers are no longer grandfathered under this counterfactual policy simulation, those boilers now face an implicit emissions tax. Many of these older boilers now choose to adjust their capital to a new vintage to reduce their emissions, especially since they have delayed productivity-enhancing upgrades to avoid that tax. In the next year, many more boilers are newer and cleaner. This can also be seen in the "No Grand." curve in Figure 4: once the law is changed, the rate of investment is much higher due to the shock in the policy. Though the number scrapping quickly gets closer to the baseline levels, the initial increase in brand new boilers reduces emissions throughout the simulation period. The magnitude of the change in both emissions and investment is large; in the first year after the policy change 350 additional boilers scrap, and by 1995 emissions are less than 40% of baseline levels. These large outcomes are from a very large policy change: since 87% of boilers are initially grandfathered, eliminating grandfathering amounts to a massive policy shift. The result I find is slightly smaller in magnitude than that of [10], who find that if all plants were subject to NSR standards (that is, if grandfathering were eliminated), then emissions of SO_2 and NO_X would fall by 75%. Here, only SO_2 emissions policy is modeled.

The results from the "Tax Cut" and "Tax Hike" counterfactuals follow intuition. When the implicit tax on emissions intensity is increased by 10%, emissions fall, and ten years after the policy change emissions are 10% lower than the baseline. A 10% decrease in the value of the implicit tax sees a rise in emissions, though of a smaller magnitude. Emissions in 1995 are only 4% higher than in the baseline.

The effects of the counterfactual simulations on emissions can be compared with the scrapping simulation results in Figure 4. Under the "Tax Cut" scenario, a larger number of boilers initially scrap and replace, while under the "Tax Hike" scenario, a smaller number do. Thus in the "Tax Hike" scenario, boilers are older on average than in the baseline scenario. Since older boilers are on average dirtier, one would expect that in the presence of an increase in the virtual emissions tax, boilers would scrap at a higher rate to avoid paying this increased tax, and likewise after a decrease in the implicit tax rate fewer boilers would scrap. However, here the effects of grandfathering come into play: grandfathered boilers have a valuable status that is lost when they scrap and replace. Increasing the tax rate on emissions increases the value of this status, and makes grandfathered boilers less likely to scrap than in the baseline simulation.

This effect of grandfathering on scrapping rates can be seen in Figure 5. This presents the simulation results for scrapping for the "Tax Cut" and "Tax Hike" counterfactuals, but separately for boilers that are initially grandfathered and for those that are initially non-grandfathered. The top panel of Figure 5, which shows the results for boilers that are initially grandfathered, clearly shows the results of changes in the virtual tax rate on scrapping outcomes. When the tax is increased, fewer of these grandfathered boilers scrap, since they have even more incentive to hold on to their valuable grandfathered status. On the other hand, the bottom panel of Figure 5, which shows the results for boilers that are initially non-grandfathered, shows no such behavior. After a tax increase, more boilers choose to scrap to newer, cleaner plants, to avoid paying the higher tax on emissions intensity. On the other hand, after a tax decrease, the scrapping rate is slightly lower. On the whole, as shown in Figure 4, the effect from grandfathered boilers dominates the effect from non-grandfathered boilers, in large part because such a high fraction of boilers (87%) are grandfathered. Figure 5 also demonstrates how grandfathering creates a schism in the reaction of boilers to policy changes, with grandfathered and non-grandfathered boilers responding in opposite directions. The reduction in new investment comes only from grandfathered boilers, and hence the magnitude and even the existence of a reduction in investment overall depends on the fraction of the fleet in the cross-sectional distribution that is grandfathered at the onset of the policy change.

Comparing the results for emissions in Figure 3 with those for scrapping in Figures 4 and 8 presents a puzzle. No evidence of a perverse result for emissions is found, yet a perverse result for scrapping is found. These results do not conform to some previous theoretical and empirical findings of a perverse effect of grandfathered policies, including the results in [1] for automobile emissions standards.[29] Though inconsistent with [1]'s results for automobiles, the results here that do not find a perverse effect are consistent with results from [2], who also study electric power plants. They find that in the absence of regulations, emissions would have increased by 34.6%.

If an increase in the tax rate leads to less scrapping and hence older boilers than in the baseline, why isn't the resulting level of emissions from these older, dirtier boilers higher than in the baseline? Two reasons could account for this. First, the abatement choices of coal boilers could differ under the different scenarios. Changes in the implicit tax rate could affect how many coal boilers add a scrubber and how many choose to use clean coal. Second, the choice between coal and non-coal boilers could be affected by the tax rate. Since non-coal boilers emit much less SO_2, replacing scrapped boilers with non-coal boilers instead of coal boilers reduces emissions, even with no difference in the overall scrapping rate.

The second reason, not the first, explains the results here. The changes in abatement activity do not explain why emissions fall after a tax increase despite scrapping rates also falling. Between the three counterfactuals and the baseline, the utilization rate of scrubbers does not change. In fact, in all simulations, including the baseline, no boilers add scrubbers over the course of the simulation period.[30] The use of clean coal is different from the baseline usage in all three counterfactuals, but the differences cannot account for the observed differences in emissions. Instead, the simulated counterfactual policy scenarios affect the choice of scrapped boilers between replacing with coal and non-coal units. The top panel of Figure 6 presents the number of boilers scrapping and replacing with coal boilers, and the bottom panel presents the number scrapping and replacing with non-coal boilers. Each panel presents the absolute difference of the number scrapping from the baseline for each counterfactual. The first panel shows that the difference between the scrapping rates in the "Tax Cut" and "Tax Hike" counterfactuals is wholly accounted for by the different number of boilers replacing with coal boilers. In the "Tax Cut" counterfactual, 50 more boilers are scrapped and replaced with new coal boilers compared to the baseline, whereas in the "Tax Hike" scenario 40 fewer boilers are scrapped and replaced with new coal boilers. The difference between these two counterfactuals and the baseline in the number of boilers scrapping and replacing with non-coal boilers is almost zero. The apparent contradiction between the outcome for scrapping rates and the outcome for emissions is thus resolved: although a tax decrease results in a higher scrapping rate and hence younger boilers compared to the baseline scenario, the additional boilers that do scrap are replaced by coal rather than non-coal boilers. These boilers have a higher SO_2 emissions rate, and hence emissions decreases after the tax decrease relative to the baseline. The opposite holds for a tax increase. This explanation also likely accounts for the results from prior literature, here explained through the dynamic model.

The results for the "No Grand." counterfactual simulation show quite a different pattern. The top panel of Figure 6 shows that the number of boilers scrapping and being replaced with coal boilers is no different between the baseline and "No Grand." By contrast, the bottom panel shows that a much larger number of boilers scrap and are replaced with non-coal boilers in this counterfactual. Thus, the change in the overall scrapping rate, shown in Figure 4, is almost entirely driven by this different rate of replacement with non-coal boilers. Because these boilers emit much less SO_2 than coal boilers, emissions under this counterfactual are much lower than the baseline, as shown in Figure 3. Finally, Figure 7 summarizes the differences in types of boilers across the three counterfactual simulations by presenting the total number of coal boilers in each simulated year, as compared to the number in the baseline scenario. Because of the different scrapping and replacement rates described above, the number of coal boilers in the economy is highest under the "Tax Cut" simulation and lowest under the "No Grand." simulation. Comparing this figure to Figure 3 demonstrates that the simulated changes in emissions are largely driven by boiler type.

How realistic is the assumption that utilities are able the change the boiler type from coal to non-coal or vice versa, and by how much does this assumption affect the results? Plants are often located near a primary fuel source or access point: virtually all coal plants are located near railway lines or at the mouth of a mine, and natural gas plants have laterals that connect to major pipelines. It should be noted, though, that replacement in this model does not require that the boiler be located in the exact same location. A utility could scrap its coal plant located on a railway line and build a gas plant located near a pipeline. However, it may be that utilities have less freedom to change plant type than the model here provides them, for geographic or other constraints (perhaps a utility operates in an area with no access to any natural gas pipelines). Thus, I estimate and perform the same counterfactual simulations on a model where utilities do not have the option of changing the boiler type: coal boilers can be replaced only with coal boilers and likewise for non-coal boilers. The counterfactual results for emissions are almost identical to those of the original model: eliminating grandfathering reduces emissions by about 55% compared to the baseline. However, the mechanism by which this emissions reduction occurs is different. In the original model, emissions dropped because many coal boilers were replaced with non-coal boilers. That is not possible here; instead, many more coal boilers add scrubbers, bringing down total emissions. By the tenth year of the simulation, the number of scrubbers increases by a factor of four in the no grandfathering counterfactual compared to the baseline.

Finally, I examine how sensitive the results are to the choice of modeling the policy as an implicit tax on emissions intensity. I estimate an alternative model where the tax is levied on emissions level rather than emissions intensity. In the "Tax Cut" and "Tax Hike" counterfactual simulations of this model, the results are nearly identical to the base model. A tax cut results in more replacement of older plants and higher emissions; a tax hike has the opposite effect. In contrast, while the "No Grand." results are qualitatively similar, the magnitude of the difference between the counterfactual and the baseline simulation is much less pronounced. Removal of grandfathering results in a drop in emissions compared to the baseline, but only by about 5%, compared to the 60% drop in this counterfactual in the original model. This difference stems from the fact that utilities now have the option of meeting the pollution policy through reduced operating intensity, an option that was not possible when the policy was based on emissions intensity. In the "No Grand." counterfactual, boilers that were grandfathered and are newly subject to the implicit tax can choose to reduce operating intensity rather than have to scrap and replace. They do so, and therefore there the reduction in emissions in much less pronounced. As argued earlier, this modeling choice is a poorer representation of actual policy than the implicit tax on emissions intensity, since NSPS have always been based on intensity.

V. Conclusion

This paper investigates how environmental regulations impact plant investment decisions and plant emissions. Regulations that are grandfathered can have short-term perverse effects. When grandfathered regulations are strengthened, grandfathered units may avoid being subject to the regulation by withholding investment in newer capital. If newer capital is less polluting, the reduction in new investment may increase emissions. Also, when grandfathered regulations are weakened, emissions can decrease in the short run, as grandfathered plants lose their valuable status that keeps them from investing. For electric power plants subject to the Clean Air Act, I find that removing grandfathering from the regulations in 1985 leads to a 60% decrease in emissions by 1995. A marginal change in the stringency of grandfathered regulations leads to a perverse effect on investment but not on emissions. This is because utilities can choose boilers of different emissions intensities.

Future research could extend the model to answer other questions. Here, investment only occurs on the extensive margin: plants make a discrete choice decision about whether or not to scrap. One could allow for a continuous level of investment. The policy here is constant and without uncertainty. Adding uncertainty allows for a policy that changes over time. Alternative policies can also be considered within this model. For example, to counteract the investment disincentive created by grandfathering, regulators could provide direct subsidies to investment. Finally, the fixed cost parameters F_c, F_{nc}, and G are constant over time and firms in the model. Capital investment costs are likely to change and react to macroeconomics conditions ([11]), so these adjustment cost variables may be modeled as endogenous.

When modeling the impacts of policy that contain grandfathering provisions, whether in environmental regulations or elsewhere, it is important to consider how these provisions affect the behavior of regulated firms or individuals and the potential for unintended consequences. A dynamic structural model as presented here can capture the regulatory impacts and the incentives created by grandfathering. The cross-sectional distribution of plant ages and grandfathering statuses, the relative profitability and emissions intensity of capital of different vintage, and the ability of plants to invest in abatement measures or choose among different types of capital all can affect how regulated entities respond to these policies.

References

1 Gruenspecht, Howard. Differentiated Regulation: The Case of Auto Emissions Standards. *American Economic Review*, 81, 3 (May 1982), 328-331.

2 Nelson, Randy, Tietenberg, Tom, and Donihue, Michael. Differential Environmental Regulation: Effects on Electric Utility Capital Turnover and Emissions. *Review of Economics and Statistics*, 75, 2 (May 1993), 368-373.

3 Levinson, Arik. Grandfather Regulations, New Source Bias, and State Air Toxics Regulations. *Ecological Economics*, 28, 2 (February 1999), 299-311.

4 Maloney, Michael and Brady, Gordon. Capital Turnover and Marketable Pollution Rights. *Journal of Law and Economics*, 31, 1 (April 1988), 203-226.

5 List, John, Millimet, Daniel, and McHone, Warren. The Unintended Disincentive in the Clean Air Act. *Advances in Economic Analysis and Policy*, 4, 2 (2004), Article 2.

6 Bushnell, James and Wolfram, Catherine. *Enforcement of Vintage-Differentiated Regulations: The Case of New Source Review*. 2008.

7 Rust, John. Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher. *Econometrica*, 55, 5 (September 1987), 999-1033.

8 Keohane, Nathaniel, Mansur, Erin, and Voynuv, Andrey. Averting Regulatory Enforcement: Evidence from New Source Review. *Journal of Economics and Management Strategy*, 18, 1 (Spring 2009), 75-104.

9 Carlson, Curtis, Burtraw, Dallas, Cropper, Maureen, and Palmer, Karen. Sulfur Dioxide Control by Electric Utilities: What Are the Gains from Trade? *Journal of Political Economy*, 108, 6 (December 2000), 1292-1326.

10 Biewald, Bruce, White, David, Woolf, Tim, Ackerman, Frank, and Moomaw, William. *Grandfathering and Environmental Comparability: An Economic Analysis of Air Emission Regulations and Electricity Market Distortions*. National Association of Regulatory Utility Commissioners, 1998.

11 Wald, Matthew. Costs Surge for Building Power Plants. *The New York Times* (July 10, 2007).

12 Adda, Jerome and Cooper, Russell. Balladurette and Juppettee: A Discrete Analysis of Scrapping Subsidies. *Journal of Political Economy*, 108, 4 (August 2000), 778-806.

13 Adda, Jerome and Cooper, Russell. *Dynamic Economics: Quantitative Methods and Applications*. MIT Press, Cambridge, 2003.

14 Berry, Steven, Levinsohn, James, and Pakes, Ariel. Voluntary Export Restraints on Automobiles: Evaluating a Trade Policy. *American Economic Review*, 89, 3 (June 1999), 400-430.

15 Bovenberg, Lans and de Mooij, Ruud. Environmental Levies and Distortionary Taxation. *American Economic Review*, 84, 4 (September 1994), 1085-1089.

16 Cooper, Russell, Haltiwanger, John, and Power, Laura. Machine Replacement and the Business Cycle: Lumps and Bumps. *American Economic Review*, 89, 4 (September 1999), 921-946.

17 Doms, Mark and Dunne, Timothy. Capital Adjustment Patterns in Manufacturing Plants. *Review of Economic Dynamics*, 1, 2 (April 1998), 409-429.

18 Fullerton, Don and Heutel, Garth. The General Equilibrium Incidence of Environmental Taxes. *Journal of Public Economics*, 91, 3-4 (April 2007), 571-591.

19 Geman, Ben. Energy Law's Emissions Mandate Largely Exempts Corn Ethanol. *Greenwire* (March 28, 2008).

20 Gentzkow, Matthew. Valuing New Goods in a Model with Complementarity: Online Newspapers. *American Economic Review*, 97, 3 (June 2007), 713-744.

21 Gourieroux, Christian and Monfort, Alain. *Simulation-Based Econometric Models*. Oxford University Press, Oxford, 1996.

22 Joskow, Paul. Contract Duration and Relationship-Specific Investments: Empirical Evidence from Coal Markets. *American Economic Review*, 77, 1 (March 1987), 168-185.

23 Lee, Bong-Soo and Ingram, Beth Fisher. Simulation Estimation of Time-Series Models. *Journal of Econometrics*, 47, 2-3 (February-March 1991), 197-205.

24 McFadden, Daniel. A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration. *Econometrica*, 57, 5 (September 1989), 995-1026.

25 Metcalf, Gilbert. Environmental Levies and Distortionary Taxation: Pigou, Taxation and

Pollution. *Journal of Public Economics*, 87, 2 (February 2003), 313-322.

26 Nilsen, Oivind and Schiantarelli, Fabio. Zeros and Lumps in Invenstment: Empirical Evidence on Irreversibilities and Nonconvexities. *Review of Economics and Statistics*, 85, 4 (November 2003), 1021-1037.

27 OFFICE OF TECHNOLOGY ASSESSMENT. *Acid Rain and Transported Air Pollution: Implications for Public Policy*. Washington DC, 1984.

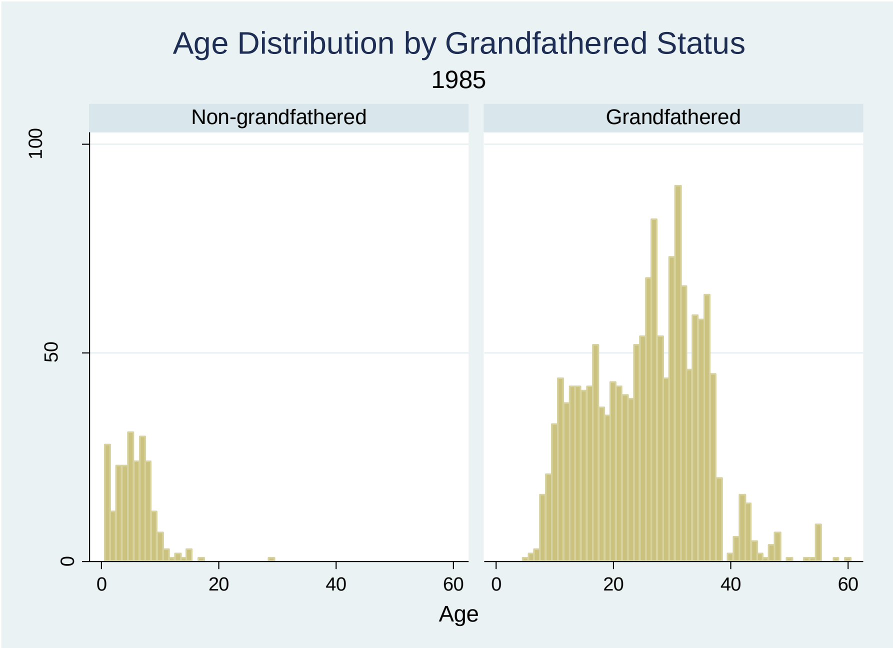
28 Shavell, Steven. On Optimal Legal Change, Past Behavior, and Grandfathering. *Journal of Legal Studies*, 37, 1 (January 2008), 37-85.

29 Stavins, Robert. Vintage Differentiated Environmental Regulation. *Stanford Environmental Law Journal*, 25, 1 (2006), 29-66.

30 Ryan, Stephen. *The Costs of Environmental Regulation in a Concentrated Industry*. Massachusetts Institute of Technology Department of Economics, 2009.

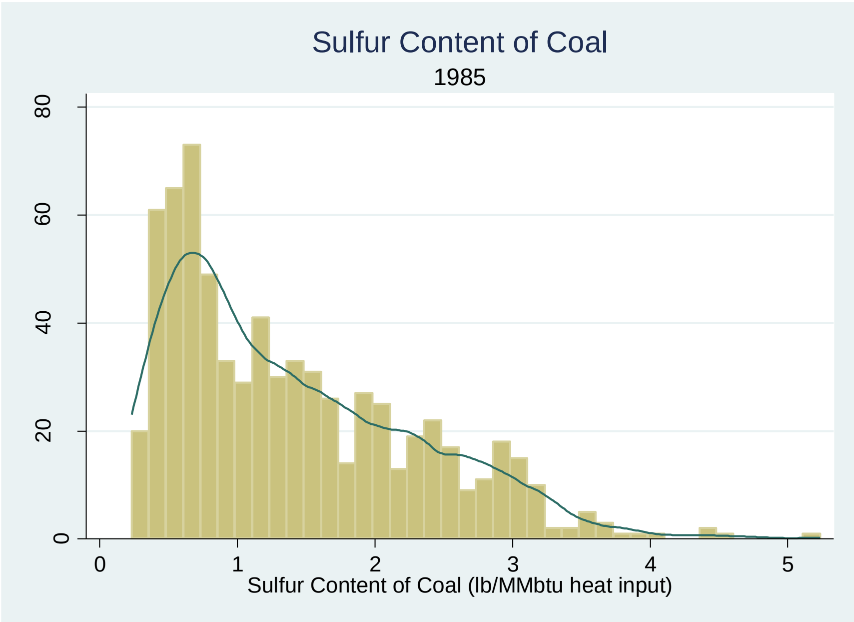
31 Lange, Ian and Bellas, Allen. The 1990 Clean Air Act and the Implicit Price of Sulfur in Coal. *The B.E. Journal of Economic Analysis & Policy (Topics)*, 7, 1 (2007).

Figure 1



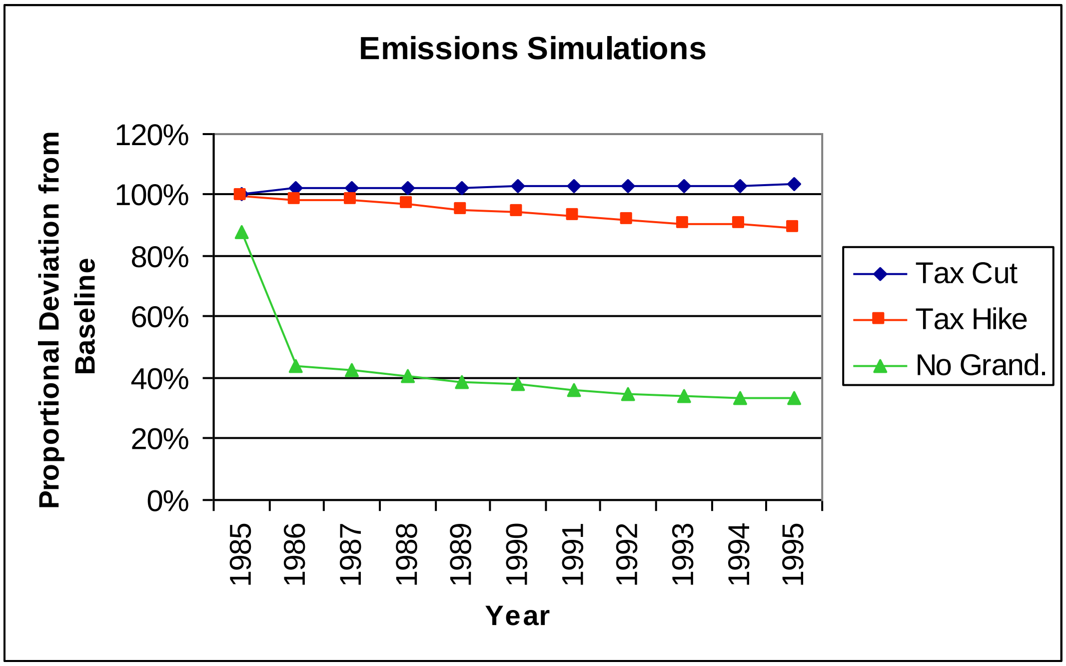
Notes: Data source is EIA-767, 1985. The *x*- axis is the age of the boiler; the *y*-axis is the number of boilers of that age. Grandfathered boilers are those not subject to any New Source Performance Standard.

Figure 2



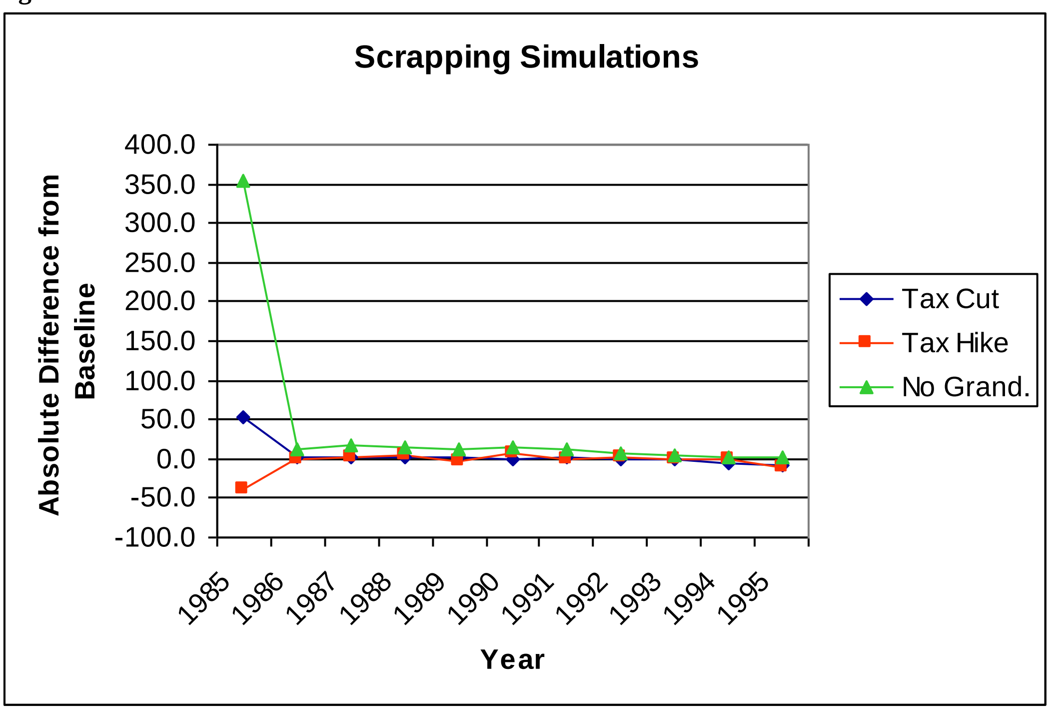
Notes: Data source is EIA-767, 1985. The curve is a nonparametric density estimation based on an Epanechnikov kernel and optimal bandwidth.

Figure 3



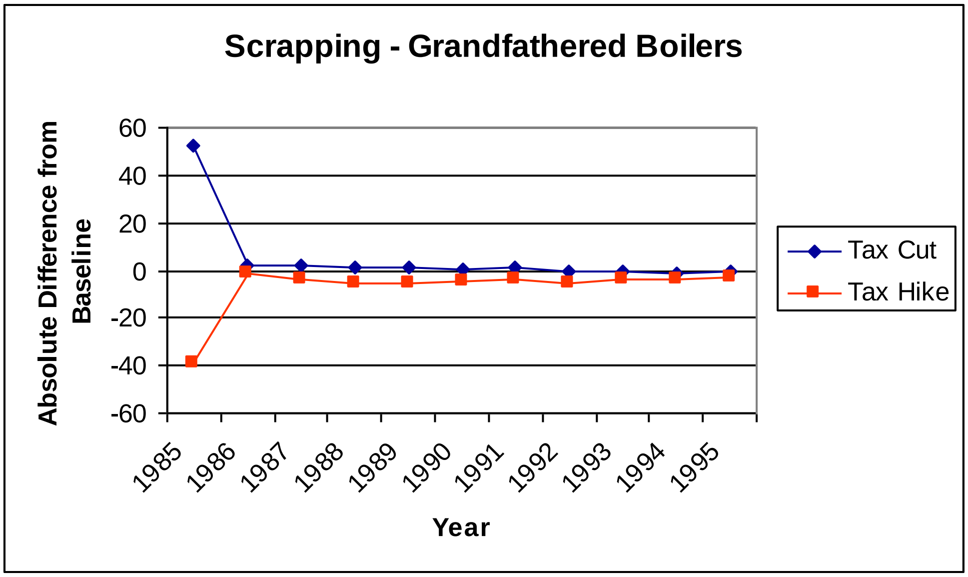
Notes: Results are from counterfactual simulations as described in the text. Values presented are proportional deviations from baseline simulation.

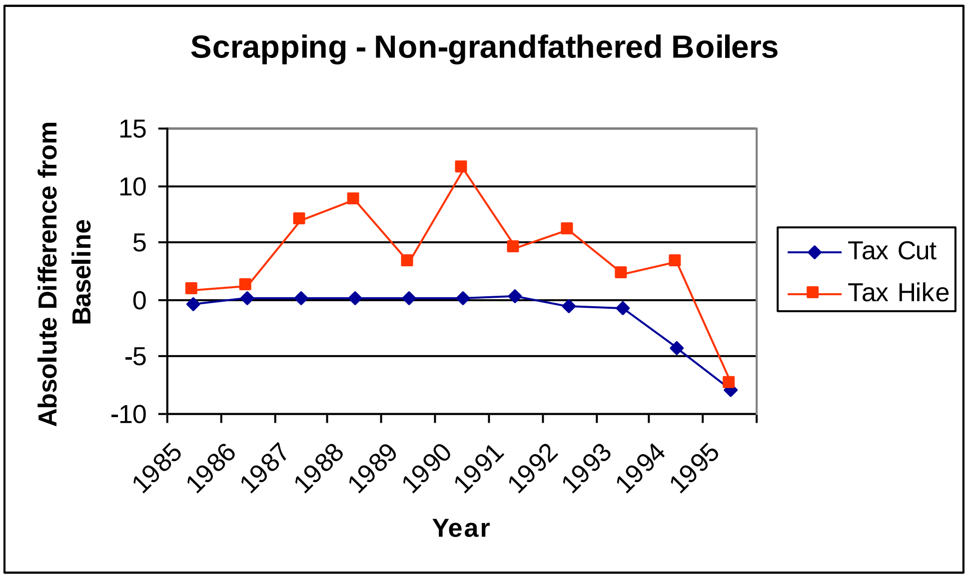
Figure 4



Notes: Results are from counterfactual simulations as described in the text. Values presented are absolute deviations from baseline simulation.

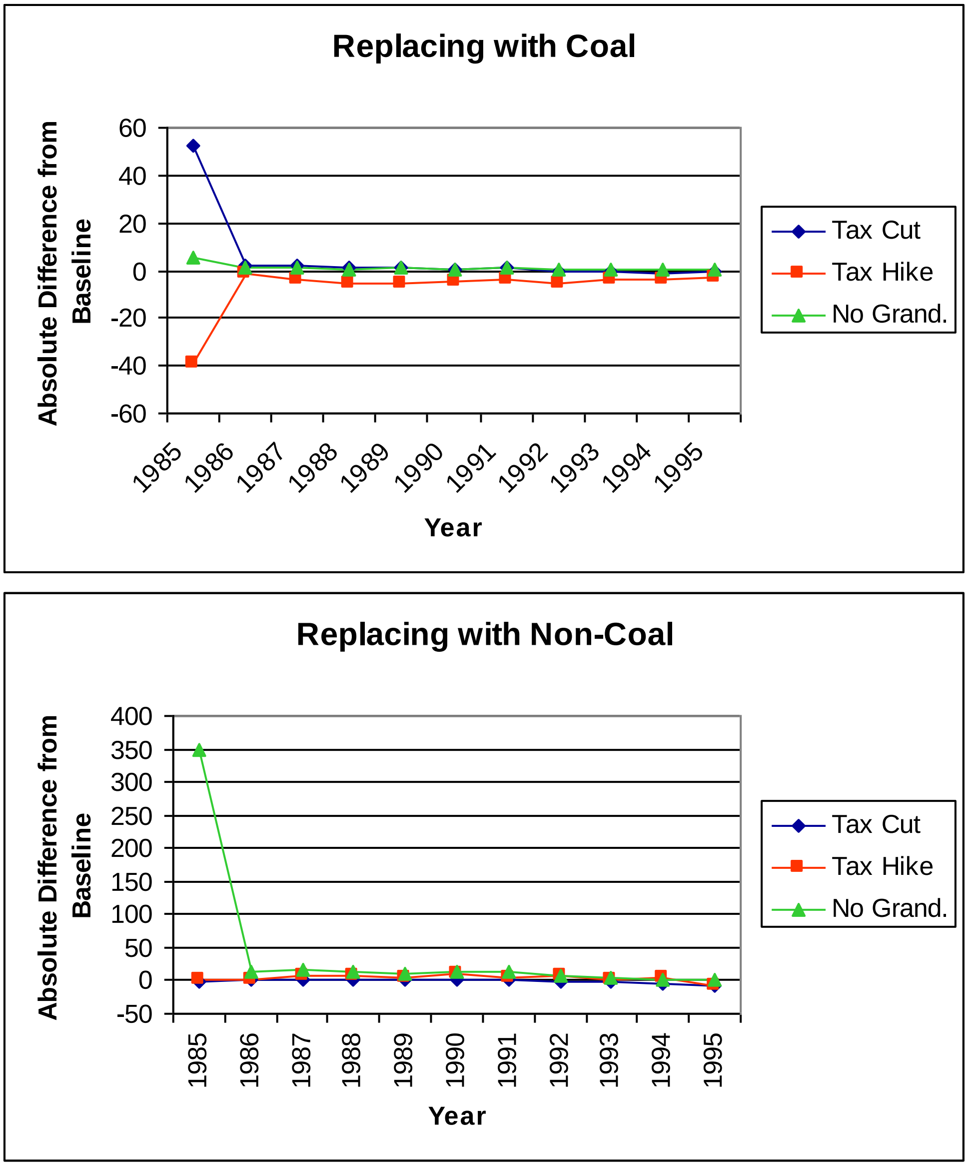
Figure 5





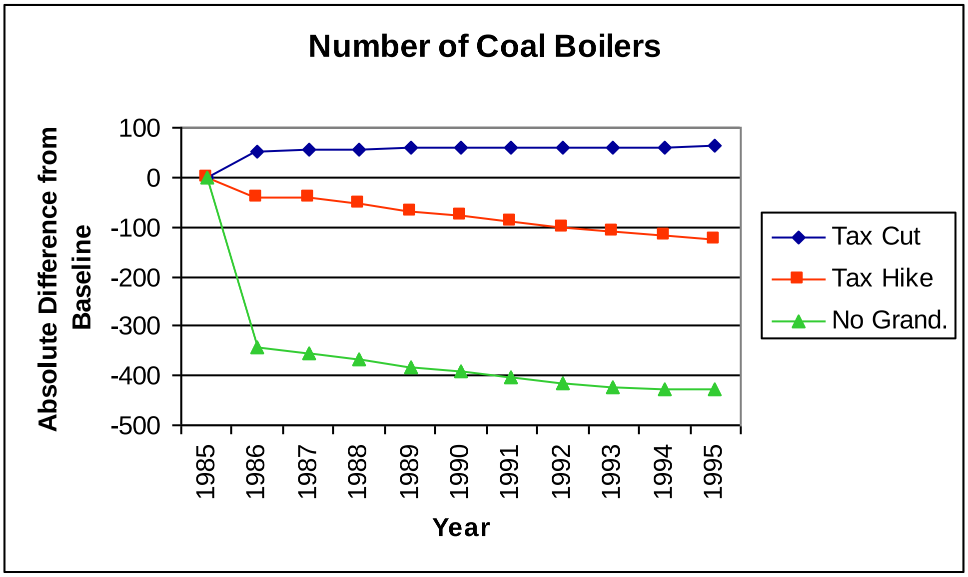
Notes: Results are from counterfactual simulations as described in the text. Values presented are absolute deviations from baseline simulation. The top panel presents simulations results for boilers that are initially grandfathered. The bottom panel presents the results for boilers that are initially non-grandfathered.

Figure 6



Notes: Results are from counterfactual simulations as described in the text. Values presented are absolute deviations from baseline simulation. Top panel presents the number of boilers scrapping and replacing with coal boilers; the bottom panel presents the number scrapping and replacing with non-coal boilers.

Figure 7



Notes: Results are from counterfactual simulations as described in the text. Values presented are absolute deviations from baseline simulation.

Table 1

Summary of Variables in Dynamic Model		
State Variables	Choice variables	Transition process
v – Boiler age, integer	z – Replacement decision, discrete (z = 0 if no replacement, z = 1 if replaced with a coal boiler, z = 2 if replaced with non-coal boiler)	$v' = v + 1$ if $z = 0$ $v' = 1$ if $z > 0$
b – Boiler type, binary (b = 1 for coal, b = 0 for non-coal)	x – Build scrubber decision, binary	$b' = b$ if $z = 0$ $b' = 1$ if $z = 1$ $b' = 0$ if $z = 2$
s – Scrubber present, binary	c – Use clean coal decision, binary	$s' = 1$ if ($s = 1$ and $z = 0$) or ($x = 1$) $s' = 0$ otherwise
u – Grandfathered status, binary (u = 0 if grandfathered, u = 1 if not grandfathered)	m – Intensity of use decision, continuous	$u' = u$ if $z = 0$ $u' = 1$ if $z > 0$
p – Price differential between high- and low-sulfur coal, continuous		p' determined exogenously
A – productivity shock, continuous		A' determined exogenously
B – emissions shock, continuous		B' determined exogenously

Table 2

Data Moments Used in MSM Estimation					
Annual Moments					
	1985	1986	1987	1988	1989
Fraction grandfathered	.8736018	.8659091	.8609502	.859366	.8571429
Fraction coal boilers	.5755079	.5869809	.6020702	.604611	.6030006
Fraction of coal with scrubber	.1354962	.1453155	.1451767	.1458532	.1492823
Fraction of coal using clean coal	.3435115	.3518164	.3906399	.4566254	.3617225
	1990	1991	1992	1993	1994
Fraction grandfathered	.8564946	.8510029	.8485023	.8463768	.8435092
Fraction coal boilers	.5996563	.6087963	.6101104	.6153396	.6230586
Fraction of coal with scrubber	.1509074	.1625475	.1685714	.1712655	.1783317
Fraction of coal using clean coal	.3352436	.3612167	.3314286	.3644148	.4074784
Categorical Moments					
	Age <= 15	15 < Age <= 25 and grandfathered	15 < Age <= 25 and not grandfathered	25 < Age <= 35	Age > 35
Fraction scrapping	.0052647	.0091001	.0233645	.0233645	.0705286
Fraction of coal with scrubber	.4264165	.1077819	.5617978	.0556046	.0382166
Fraction of coal using clean coal	.5283297	.3499792	.511236	.3415711	.2845011
Fraction coal boilers	.7680136	.5695662	.5615142	.5695375	.5675839
Fraction in category	.1630	.1927	.0144	.2726	.3573

Notes: Data source is EIA-767, 1985-1995.

Table 3

Emissions Function Estimation Results									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Coal					Non-coal			
	lnemis	lnemisrate	lnemisrate	lnemisrate	lnemisrate	lnemis	lnemisrate	lnemisrate	lnemisrate
lnage	0.128***	0.138***	0.165***	0.0504***	0.297***	0.166	-0.189	0.438**	0.0258
	(0.0102)	(0.00850)	(0.0102)	(0.0166)	(0.0239)	(0.148)	(0.136)	(0.190)	(0.141)
scrubber	-1.308***	-1.306***	-1.239***	-1.213***	-1.458***				
	(0.0328)	(0.0330)	(0.0373)	(0.0622)	(0.0384)				
lsc	-1.059***	-1.058***	-0.997***	-0.907***	-0.924***				
	(0.00961)	(0.00958)	(0.0120)	(0.0129)	(0.0169)				
lsc_scrubber	0.285***	0.283***	0.130***	0.385***	0.680***				
	(0.0435)	(0.0435)	(0.0494)	(0.0826)	(0.0537)				
lnmmbtus	0.994***					1.207***			
	(0.00404)					(0.0308)			
nonattainment			-.0964***					-0.697	
			(0.0190)					(0.495)	
Coal-gas ratio			-0.129					-0.236	
			(0.0972)					(1.176)	
Coal-oil ratio			-0.0700					-0.125	
			(0.0597)					(1.824)	
Capacity				-1.7e-7***					.000226***
				(3.03e-08)					(4.87e-05)
Heat Rate				5.19e-07					-3.39e-7***
				(5.25e-07)					(9.74e-08)
Hours				-1.0e-5***					6.5e-5***
				(2.44e-06)					(1.64e-05)
Constant	-6.697***	-6.832***	-6.791***	-6.592***	-7.720***	-17.03***	-13.57***	-14.98***	-14.45***
	(0.0892)	(0.0333)	(0.0578)	(0.0665)	(0.0930)	(0.837)	(0.602)	(1.023)	(0.624)
Observations	10837	10837	6530	6209	5154	2940	2940	1446	2940
R-squared	0.937	0.806	0.819	0.722	0.734	0.719	0.646	0.579	0.650

Notes: Coal emissions data in columns 1-4 are calculated based on EIA-767, 1985-1995. Coal emissions data in column 5 and non-coal emissions data in columns 6-9 are from eGrid, 1996-2000. Columns 3 and 8 include only boilers in states that had not begun electricity market restructuring by 2000. The dependent variable is the log of tons of SO_2 emissions or the log of tons of SO_2 emissions per MMbtus heat input. All columns include unreported state and year fixed effects. Robust standard errors are in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 4

Revenue Function Estimation Results								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Non-coal				Coal			
Age	-1.8e+06***	-891911***	-920688***	-777489***	-9.31e+6***	-6.74e+6***	-6.72e+6***	-3.59e+6***
	(256324)	(223868)	(233285)	(130785)	(496721)	(1.599e+06)	(1.603e+06)	(226540)
Coal-gas ratio			7.018e+06				1.502e+07	
			(7.495e+06)				(1.204e+07)	
Coal-oil ratio			-9.773e+06				-2.10e+7***	
			(1.044e+07)				(6.154e+06)	
p·c					-124463	-95069***	-92138**	-43314
					(115035)	(36650)	(36725)	(35468)
Constant	9.31e+7***	7.07e+7***	7.61e+7***	2.657e+07	4.91e+8***	3.80e+8***	3.87e+8***	3.30e+8***
	(8.937e+06)	(7.240e+06)	(7.926e+06)	(5.880e+07)	(1.442e+07)	(2.724e+07)	(2.756e+07)	(1.040e+08)
Observations	3096	3096	2907	3096	3760	3760	3760	3760
R-squared	0.347	0.334	0.322	0.331	0.527	0.510	0.510	0.419
Number of boilers	--	720	704	720	--	975	975	975
Estimation	OLS	RE	RE	RE	OLS	RE	RE	RE
Error structure	Clustered at Boiler	Clustered at Boiler	Clustered at Boiler	AR(1)	Clustered at Boiler	Clustered at Boiler	Clustered at Boiler	AR(1)

Notes: Data source is EIA 767, 1985-1995. Relative fuel price ratio data are from EIA-423, 1985-1995. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table 5

Structural Parameter Estimation Results	
F_c	1803.34 (1.61)
F_{nc}	84.95 (0.60)
G	61.41 (0.45)
τ	2989.77 (258.62)
σ_A^2	0.1104 (0.4851)
P	0.8727 (0.1207)

Notes: Standard errors are in parentheses. Estimates come from simulated method of moments, matching moments from Table 2.

Appendix: Not for Publication

Solving the model with Value Function Iteration

The discrete choice model presented in Section 1 does not have an analytic solution in its most general form. It is thus solved numerically using value function iteration (VFI). I briefly described how this technique is applied. Code implementing this and further documentation are available on the author's website.

For a given functional form and set of parameter values, VFI can find the value function and policy functions of the model. Suppose that the functional forms and parameters are known. The first step in implementing the solution method is to choose a discrete state space for all state and choice variables. The boiler's age, v, is in years and hence already discrete. I choose 90 years old as the maximum age possible, implying that when a boiler passes age 90 it must be scrapped (the oldest operating boiler in the data set is 89 years old). A boiler's scrubbed status s, grandfathered status u, and type (coal or non-coal) b are all binary variables. The clean-dirty coal price differential p is continuous and hence must be discretized. I divide the entire distribution of values of p into ten deciles, and assign each boiler to one of these deciles. A boiler's value of p does not change over time.

The final two state variables are the productivity shock and the emissions shock A and B. The first shock, A, is discretized into ten separate values. The distribution of A is lognormal with a median of one and a variance of σ_A^2. This distribution is divided into ten deciles, and the discretized value within each decile is the median within that decile. The emissions shock B is not estimated here, since the emissions function comes from an engineering equation rather than estimated emissions data. Thus, this shock is omitted from the analysis (or it takes only one value, unity). The Markov process that governs the evolution of the shock must also be parameterized.

Three of the choice variables are discrete. The scrapping choice z can be 0 (no scrapping), 1 (replacing with a coal plant) or 2 (replacing with a non-coal plant). Coal plants face a binary decision to add a scrubber or not (x) and a binary decision to use clean or dirty coal (c). The decision over operating intensity m is continuous. However, this decision is estimated in an earlier stage of the process and thus does not need to be solved through VFI. For each state, there is an optimal level of operating intensity m that is taken as a given. All other decisions are made under the assumption that the boiler is optimizing over this dimension.

The value function $V(v,s,u,b,p,A,B)$ must thus be solved for each of its 72,000 unique values (90*2*2*2*10*10*1). I begin with an initial guess of the value function (V_0), say, that it equals one everywhere. For each state, the optimal policy can be found as the policy that maximizes the sum of the single period profit function and the discounted value of the appropriate value function given V_0. The optimal choice gives a new value function V_1. The procedure is repeated, now with V_1 taking the place of V_0 and giving a newer resulting value function. This is iterated N times until the difference between V_N and V_{N+1} is not too great. The actual value of the difference between these two values that triggers an end to the iterations is somewhat arbitrary. When this difference is defined as simply the mean value of the pairwise differences in all dimensions of V, I find that a value of 0.001, compared to an average value function of around 65, to be a reasonable value. Reducing that cutoff value increases the computational time, but does not bring about substantive changes in the model's outcomes.

Running the VFI routine on a Windows PC with a 2.6 GHz processor and 512MB of memory takes about ten minutes to complete.

Method of Simulated Moments estimation routine

The intuition behind this application of MSM is as follows. From the data, I create a vector of moments ω. I choose a set of parameters Θ and solve the model using value function iteration. With the model solved, I simulate data and create a simulated set of those moments $\omega_s(\Theta)$. I repeat the simulation S times. The final estimate of Θ is the set of parameters that minimizes $[\omega - \frac{1}{S}\sum_s \omega_s(\Theta)]W[\omega - \frac{1}{S}\sum_s \omega_s(\Theta)]'$, where W is a weighting matrix. The weighting matrix used is an estimate of the optimal weight matrix. Given this weight matrix, the variance of the parameter estimate is given by $(1+\frac{1}{S})[\frac{\partial \omega'}{\partial \Theta} W^{-1} \frac{\partial \omega}{\partial \Theta'}]^{-1}$, which can be approximated numerically.[31]

The weight matrix that results in the optimal estimator is the covariance matrix of the difference between the true moments and the estimated moments. This optimal weighting matrix W^* can be estimated by

$$W^* = \frac{1}{N}\sum_{i=1}^{N}[\omega_i - \frac{1}{S}\omega_{is}]\cdot[\omega_i - \frac{1}{S}\omega_{is}]' + \frac{1}{S}\frac{1}{N}\sum_{s=1}^{S}\sum_{i=1}^{N}[\omega_{is} - \frac{1}{L}\omega_{il}]\cdot[\omega_{is} - \frac{1}{L}\omega_{il}]' .$$

Here, ω_i is shorthand for the data moment calculated from observation i, ω_{is} is the moment from the simulation s of that moment, and ω_{il} is the moment from an additional set of L simulations.[32] The simulated moments are taken from a first stage estimate of Θ, where the identity matrix is used as a weight matrix. This is a consistent estimate of the optimal weight matrix as $L \to \infty$.

Note that as the number of simulations, S, increases, the variance of the estimator decreases. I use $S = 20$ simulations. Furthermore, the derivative matrix $d\omega/d\theta$ is approximated by looking at small but discrete changes in parameter values about the estimated values, and observing the resulting changes in the simulated moments.

Implementing MSM thus involves four steps. First, I find the parameter values that minimize the objective function when using the identity matrix for the weight matrix. Second, using those first-stage estimates I estimate the optimal weight matrix. Third, I re-run the optimization routine using the optimal weight matrix (though the resulting estimates are more efficient, the point estimates of the parameter values are barely affected). Fourth, I calculate the variance of the estimates by approximating the derivative of the moments with respect to the parameters.

To find parameters values that minimize the objective function, I use the Nelder-Mead simplex, or "amoeba", method.[33] This routine is well-suited for searching over discontinuous functions, as is the case here with the discrete choice problem, since it does not depend on calculating gradients. However, it is still possible for this optimization routine to become stuck at a local minimum or on a flat portion of the objective function. Thus, I begin with a simple, manual grid search over a subset of parameters (keeping the rest constant) to find appropriate starting values for the simplex search. It is especially important in this application to ensure that the optimization routine is not stuck at a local minimum. Because the frequency of scrapping is so low (see the moments in Table 2), parameter values that make it never optimal for boilers to scrap result in simulated moments that are rather close to the actual moments. Small changes about these parameter values lead to no change in simulated outcomes, rendering the algorithm incapable of escaping these values. An "eyeball test" of the resulting parameters and moments is thus important.

Depending on the starting values used, implementing this MSM routine takes approximately 300 function evaluations, that is, 300 runnings of the VFI routine described above. On a Windows PC with a 2.6 GHz processor and 512MB of memory, this can take up to ten hours of computational time.[34]

Comparison of Simulated and Data Moments.

In the Appendix Table A1, I repeat from Table 2 the moments from the data and compare them to the moments that are created by simulating the economy using the estimated parameters, to see how well the model does at matching the moments. Table A1 is organized just as Table 2, with the top panel presenting the four annual moments for each year and the bottom panel the five categorical moments for each category of boilers. Within each cell, the top row is the moment evaluated from the data, replicating that in Table 2. The second row is the simulated moment evaluated from the model run using the point estimates of the parameters. In the bottom row are the 95% confidence intervals for those simulated moments. These confidence intervals are evaluated by taking 100 draws of parameter values from the estimated distribution, solving the model and simulating the moments for each draw, and inferring the resulting distribution of each moment.[35] The annual moments for the initial year, 1985, are not simulated, since those are used as starting values and are identical for all simulations and all parameter values.

The simulated annual moments are matched quite closely with the data moments. The simulated fraction of boilers grandfathered and fraction of boilers that are coal are both within one percentage point of the data moment in each year. The slight increase in the fraction of coal boilers using scrubbers is not picked up in the simulation, but by the final year the simulated moment is only 5 percentage points lower than the actual. The fraction of coal boilers using clean coal is well matched in the simulations, with the exception of 1988, in which there is an increase in clean coal use that is not captured in the simulation.

The simulated values of the categorical moments capture the overall patterns in the data moments but do not match them as closely as the matches in the annual moments. In the simulations no boilers younger than 15 years old scrap. This is actually not too far off from the data value of one-half of one percent, which represents only a small absolute number of boilers in that category scrapping. The simulated scrapping rate increases in the next age category, and the scrapping rate is higher for non-grandfathered boilers than it is for grandfathered boilers, as it is in the data. For boilers older than 25 years old, the data shows higher scrapping rates, while the scrapping rates in the simulations taper off. The simulated fraction of coal boilers using a scrubber matches the overall patterns in the data, though too few of the youngest boilers utilize scrubbers. The simulated fraction of coal boilers using clean coal is too low in the last three categories. Finally, the last two categorical moments, the fraction of coal boilers in each category and the fraction of all boilers in each category, are matched quite closely.

The 95% confidence intervals for the simulated moments tend to be smaller for the annual moments than for the categorical moments. An exception is the annual moment of fraction of coal boilers using clean coal, where the confidence interval hovers around [.23,.45]. The confidence interval for this same statistic in the categorical moments is even larger, with one interval of almost 35 percentage points. A large confidence interval, though, indicates that the moments are doing a good job identifying the parameters, since small changes in parameters lead to large changes in simulated moments. Overall, comparing the simulated with the data moments suggests that the model does a good job capturing the patterns in the data of utility decisions over abatement and scrapping. It appears that in the simulations slightly too few coal boilers use scrubbers, and too few older boilers are scrapping.

Appendix Table A1

Comparison of Data and Simulated Moments					
Annual Moments					
	1985	1986	1987	1988	1989
Fraction grand-fathered	.8736	.8659 .8508 [.8190, .8857]	.8610 .8501 [.8175, .8855]	.8594 .8479 [.8151, .8844]	.8571 .8450 [.8127, .8831]
Fraction coal boilers	.5755	.5870 .6040 [.5689, .6352]	.6021 .6047 [.5689, .6365]	.6046 .6070 [.5700, .6388]	.6030 .6098 [.5712, .6411]
Fraction of coal with scrubber	.1355	.1453 .1355 [.1250, .1770]	.1452 .1353 [.1254, .1849]	.1459 .1348 [.1242, .1932]	.1493 .1342 [.1249, .1984]
Fraction of coal using clean coal	.3435	.3518 .3376 [.2254, .4320]	.3906 .3384 [.2260, .4338]	.4566 .3408 [.2272, .4366]	.3617 .3439 [.2285, .4395]
	1990	1991	1992	1993	1994
Fraction grand-fathered	.8565 .8422 [.8104, .8820]	.8510 .8398 [.8081, .8808]	.8485 .8375 [.8057, .8800]	.8464 .8346 [.8033, .8787]	.8435 .8326 [.8013, .8778]
Fraction coal boilers	.5997 .6127 [0.5721, 0.6435]	.6088 .6151 [0.5734, 0.6456]	.6101 .6174 [0.5741, 0.6479]	.6153 .6200 [0.5751, 0.6502]	.6231 .6215 [0.5755, 0.652]
Fraction of coal with scrubber	.1509 .1336 [.1234, .2057]	.1625 .1330 [.1237, .2092]	.1686 .1325 [.1232, .2109]	.1713 .1320 [0.1232, 0.2116]	.1783 .1317 [0.1234, 0.2108]
Fraction of coal using clean coal	.3352 .3469 [0.2297, 0.4422]	.3612 .3495 [0.231, 0.4449]	.3314 .3519 [.2320, .4474]	.3644 .3551 [.2333, .4503]	.4075 .3575 [.2343, .4528]

Categorical Moments					
	Age <= 15	15 < Age <= 25 and grand-fathered	15 < Age <= 25 and not grand-fathered	25 < Age <= 35	Age > 35
Fraction scrapping	.005265 .0000 [.0000, .004300]	.009100 .005600 [.0000, .01100]	.02336 .02650 [.0000, .03540]	.02336 .007000 [.0000, .01280]	.07053 .003700 [.0001000, .005900]
Fraction of coal with scrubber	.4264 .3211 [.2832, .5636]	.1079 .07430 [.07430, .07430]	.5618 .5775 [.5528, .9703]	.05560 .04910 [.04910, .04910]	.03822 .05210 [.05210, .05210]
Fraction of coal using clean coal	.5283 .5695 [.3622, .7327]	.3500 .3511 [.2420, .4449]	.5112 .4033 [.2762, .5399]	.3416 .2470 [.1703, .3130]	.2845 .2000 [.1379, .2534]
Fraction coal boilers	.7680 .7738 [.7268, .8035]	.5696 .5712 [.5581, .5863]	.5615 .4985 [.4647, .5221]	.5695 .5754 [.5421, .5996]	.5676 .5563 [.5282, .5793]
Fraction in category	.1630 .2061 [.1706, .2349]	.1927 .2206 [.2149, .2256]	.01440 .01240 [.0114, .0129]	.2726 .3038 [.2911, .3216]	.3573 .2572 [.2466, .2704]

Notes: In each cell, the top row is the moment from the data (replicated from Table 2). The middle row is the corresponding simulated moment using the point estimates of the structural parameters. The bottom row is the 95% confidence interval of the simulated moments, generated from 100 draws from the estimated distribution of the parameters.

Endnotes

[1] The US energy bill passed in 2007 contains an ethanol mandate with a requirement that the ethanol produced over its lifetime achieves at least a 20 percent reduction in greenhouse gases compared to gasoline. However, any biorefinery built or under construction before the bill's signing is exempt from that requirement; this could include up to 13.7 billion gallons of production capacity out of the 15 billion gallons of corn ethanol envisioned by the bill [19].

[2] Technically, the CAA was passed in 1963 and amended in 1970, but the 1970 amendments are often referred to as the "Clean Air Act of 1970" because they contained the bulk of the regulations.

[3] [28] discusses when it might be socially optimal to grandfather, and [29] discusses the effects of grandfathering of environmental policy in many contexts.

[4] The data set includes fossil fuel fired steam electric power plants powered by coal, oil, and natural gas. I lump together oil and natural gas because their emissions of sulfur dioxide are negligible compared to coal plants; coal accounts for 97% of SO_2 emissions from this group.

[5] [31] finds evidence of a discrete jump in coal prices right at the CAA-imposed standard of 1.2 lbs/MMBtu.

[6] This choice is assumed to be binary rather than continuous. [17] and [26] find that investment is mostly "lumpy." Entry and exit are not modeled. The vast majority of the electric utilities are present in every year, indicating that the utility-level industry makeup is roughly constant and entry and exit are not prevalent in this industry. Capacity growth, on the other hand, is prevalent, and it can be accommodated in the model in the same way as can technological growth. See the discussion later in this section.

[7] The cost differential of switching to low-sulfur coal is modeled as a variable cost. It is possible to also include a fixed cost that is incurred when first switching coal types. Different types of coal can have different ash contents, different difficulties of pulverization, and other different characteristics that may require capital or fixed costs to accommodate the boiler. However, those capital costs are dwarfed by the differential in fuel cost ([27], p. 158).

[8] Among the many papers that model environmental policy as a tax are [18], [25], and [15].

[9] See [16] for a formalization of this intuition.

[10] This program has been studied in [9].

[11] A dynamic policy incorporating uncertainty can be modeled by through a state variable that captures the policy and evolves according to a Markov process. See [12]. If firms are responding in advance to the policy change, then the baseline estimates of the implicit emissions tax may be biased downwards.

[12] Only 12 states had begun implementing electricity market restructuring before the end of 2000, and as of September 2001 only 2.6% of total customers (residential and non-residential) were served by competitive electricity suppliers (http://www.eia.doe.gov/cneaf/electricity/chg_str/restructure.pdf).

[13] In fact, two recent papers that study the impact of NSR on existing plants, [6] and [8], both use a period of heightened enforcement that began with a series of lawsuits against utilities in November 1999 to identify this effect [8]: "Although the law [NSR] had been in place for decades, it was never vigorously enforced." (p. 1).

[14] For boilers without scrubbers, total emissions is the product of the total tons of coal, the average sulfur content of coal, and a boiler-specific emissions factor. For scrubbed boilers, the equation also considers scrubber-specific emissions reduction factors.

[15] Available online at www.epa.gov/cleanenergy/egrid.

[16] Data on emissions of nitrogen oxides (NO_X) and other pollutants are also included; I focus here solely on SO_2.

[17] Available at http://www.eia.doe.gov/cneaf/electricity/page/eia767.html for 1996-2005, but the data used here for earlier years were obtained by request from the EIA.

[18] At least three different types of NSPS can apply to the boilers, and each boiler is asked which type of NSPS it faces. While the standards differ somewhat, for simplicity I consider only a binary grandfathered policy.

[19] Available at: http://www.eia.doe.gov/cneaf/electricity/cq/cq_sum_backissues.html.

[20] For examples of this method, see [13], [21], [23], or [24]. The Appendix describes the method in detail.

[21] One can also regress emissions on a polynomial of heat input. The non-linear terms are significant but orders of magnitude smaller than the linear term.

[22] This neglects effects from long-term contracts for coal, where utilities cannot choose an optimal coal type each period since they are locked in to a contract (see, e.g., [22]). This should not present a significant distortion, since utilities are optimizing over contract terms and, on average, the outcome of that behavior should follow the outcome of choosing coal types each year.

[23] I have experimented with adding a common component to the productivity shock to capture business cycle effects, using an additional set of data moments based on business cycles. However, the common component of the shock was not found to be significant.

[24] Though β could be estimated, it may be difficult to identify due to collinearity with the fixed costs of replacement F_c and F_{nc}. (see [7], p. 1023).

[25] [30] estimates a structural model of the cement industry's response to the 1990 CAA amendments, but not grandfathering.

[26] The variance of a lognormal distribution is $\exp(\sigma^2 - 1)\times\exp(2\mu + \sigma^2)$, and here μ is zero.

[27] These are not the only counterfactual policy experiments that could be simulated. For example, one policy to counteract the effects of grandfathering is to provide subsidies to

investment. This policy might be more politically feasible then a removal of grandfathering. Alternatively, the policy could gradually phase out the grandfathering status for older plants or end it after a certain age. Note also that the no grandfathering counterfactual is similar to what occurred with the implementation of the SO_2 permit trading scheme, which did not grandfather plants.

[28] For clarity, confidence intervals are not presented but can be calculated by taking 100 draws from the asymptotic distribution of the parameters values, simulating the counterfactual for each draw, and calculating the resulting standard errors. This is the same method used to calculate the confidence intervals of the simulated moments in Table 5.

[29] Further evidence of the perverse effect of increasing emissions standards can be found in [29].

[30] This result is not too far removed from what is seen in the data, where from 1985-1994 only 36 boilers add scrubbers, a rate of 0.18%. Boilers adding scrubbers are younger than average, with a mean age of 24.3 years compared to 26.8 years for those that do not add scrubbers. Comparing the simulated moments with the data moments in Appendix Table A1, the fraction of coal boilers with scrubbers increases slightly in the data, though the simulated data do not match this pattern. Thus the simulations are slightly under-predicting the use of scrubbers, although if the level of under-prediction is the same across all simulations, then it does not affect the comparisons between the baseline and counterfactuals.

[31] Simulated maximum likelihood (SML) is another simulation method commonly used for discrete choice models. [20] finds no difference in parameter values for a model estimated with both MSM and SML.

[32] See [13], p. 89, or [21], p. 32.

[33] This is used in the "fminsearch" function in Matlab.

[34] This is not quite the time for one VFI routine (~ 10 minutes) times 300, since running the VFI routine after only a small change in parameter values takes less time than an initial running.

[35] This method is used in [14] and [20]. [14] note that it is generally superior to the delta method.