Using Cost Functions to Estimate Productivity and Abatement Efficiency: Coal Subsides and Emission Standards

Scott E. Atkinson* and Rong Luo†

Department of Economics, University of Georgia

September 26, 2018

Abstract

The recent production function estimation literature assumes that the firm maximizes profit and endogenously choose output. Although valid for many industries, researchers nearly always assume cost minimization (with exogenous output) in modeling electricity generation and other regulated industries. The derived input demand depends on output quantity. Substituting for unknown productivity in the production function implies that output is a function of itself. Instead, we assume minimization of production and pollution control costs subject to an output constraint and derive the dual cost function, which includes the parameters of the production function. We allow heterogeneity in abatement efficiency as well as generating efficiency among plants. Using 1995-2005 data on the 80 largest US coal-fired power plants, we estimate the pollution control cost function and the production function parameters with our cost-minimization approach. We examine via counterfactuals two policy initiatives of the current Administration designed to bring back coal and slow the reduction of SO_2 emissions. First, we find that a coal subsidy substantially reduces sulfur content and, ironically, SO_2 emissions. Further, an increase in emission permit allocations lowers sulfur content for plants with scrubbers, which is counterintuitive but consistent with our model. Both measures moderately reduce Btu content.

KEYWORDS: Electric power plants, production functions, cost functions, generation productivity, abatement efficiency, subsidies, emission permits

JEL CODES: D24, L11, L20, L94

*Email: `atknsn@uga.edu`

†Email: `rluo@uga.edu` The authors wish to thank participants at the 2018 International Industrial Organization Conference, the 2018 North American Chinese Economics Society, the 2018 Asian Econometric Society Meetings, and the 2018 Chinese Econometric Society Meetings for helpful comments. Rong Luo also thanks the Bonbright Center in the Department of Economics at the University of Georgia for funding support.

1 Introduction

The traditional approach of Olley and Pakes (1996) (OP), Levinsohn and Petrin (2003) (LP), Wooldridge (2009), Ackerberg, Caves, and Frazer (2015) (ACF) to production function estimation requires the assumption of maximization of profits, which implies endogenous output. Under their approaches, one derives an investment function or an input demand function, which is monotonic in unobserved productivity but does not depend output.[1] Inverting this function yields a control function which proxies for the unobserved productivity in the production function. One can estimate this production function using additional assumptions about a Markov process for productivity. Doraszelski and Jaumandreu (2013) build on this approach and derive a parametric intermediate input function. Gandhi, Navarro, and Rivers (2016) use an estimate of input elasticity to improve the identification of the traditional approach.[2]

However, profit maximization with endogenous output is not an accurate assumption for regulated US and foreign electric utility firms, regulated and deregulated electric generating plants, and regulated railroads and airlines in many countries. Firms in these industries arguably minimize costs to produce exogenous output targets determined by consumers based on prices set by regulators.[3] Even deregulated electric utilities and plants are cost-minimizing price-takers, since they belong to power-pooling organizations with many utilities and plants, like Regional Transmission Organizations (RTOs). An RTO forms a marginal cost (MC) curve using the offer price for a MW hour from each plant. The intersection of MC and aggregate demand determines the market-clearing price and quantity. Therefore, each single plant is a price-taker and will not withhold production.[4]

[1]Additional arguments are variable input prices, fixed input quantities, and output prices.

[2]A very useful summary of the earlier work is provided in Griliches and Mairesse (1998).

[3]Atkinson (2018) surveys nearly 100 cross-section and panel data studies which mainly examine regulated electricity generation (by utilities and plants), and to a lesser extent, railroads, and airlines. The vast majority of these studies estimate cost functions for plants and firms.

[4]In the absence of collusion, each offer price will equal MC and no plant will withhold production in an

If we assume minimization of variable costs but follow the OP/LP/ACF approach, we would first derive an input demand equation that is a function of productivity and output quantity.[5] Assuming invertibility of this function, a proxy function for productivity would be derived as a function of the output quantity. After substituting this proxy function into the production function, output would be as a function of itself, which yields a non-viable estimation equation.

Our first methodological contribution avoids direct estimation of the production function by formulating and estimating a variable cost function where the plant minimizes cost subject to a production function constraint. We apply our methodology to a set of exclusively coal-fired power plants and model the output-constrained minimization of the costs of production and pollution control, including sulfur abatement cost and SO_2 permit cost. The choice variables for the plants are Btu content and sulfur content of coal, which are its key characteristics. We include terms that measure heterogeneity in both generation efficiency and abatement efficiency across plants, so that the derived cost function depends on the production function parameters as well as production and abatement efficiencies. Combining the cost function with the transition functions for the efficiencies identifies the production function parameters. This approach does not require inversion of an input demand function to proxy the unobserved productivity.

Our second methodological contribution is to provide a procedure that estimates plant generation efficiency separately from pollution abatement efficiency. The literature on electric utilities has mainly focused on total factor productivity growth. Among these studies are Baltagi and Griffin (1988), Kleit and Terrell (2001), Knittel (2002), Färe, Grosskopf, Noh, and Weber (2005), Atkinson and Dorfman (2005), Bushnell, Mansur, and Saravia (2008), Atkinson and Tsionas (2016), Chan, Fell, Lange, and Li (2017), and

attempt to manipulate price. To do so would sacrifice rents which are earned only by producers.

[5] The entire set of arguments also includes variable input prices and fixed input quantities.

Atkinson, Primont, and Tsionas (2018).

We also contribute to the literature by separating plant variable costs (where captial and labor are fixed inputs) into the cost of electricity generation and pollution abatement. This allows us to model the tradeoff between the cost of coal and the cost of abatement, where the latter comprise on average about 10%, and as much as 45%, of total costs for our sample plants. To examine this tradeoff, we model the plants' endogenous choice of the sulfur and Btu content of coal. Many papers estimate a translog total variable cost function without this separation of costs, as in Christensen and Greene (1976), Gollop and Roberts (1981, 1983, 1985), and Carlson, Burtraw, Cropper, and Palmer (2000). These studies and most other papers examining electric utility or plant production only consider the choice of total Btu input or the high-sulfur/low-sulfur coal tradeoff, not the Btu-sulfur content tradeoff.[6]

Using a balanced panel of the 80 largest US coal-fired power plants from 1995-2005, we estimate a hedonic equation for the price of coal as a function of its characteristics, the abatement cost function of sulfur dioxide (SO_2), the production function parameters, and the transition functions of the efficiencies via our cost function approach. To obtain accurate implicit prices of Btu and sulfur content, we control for transportation charges by including plant-mine fixed effects with plant-mine transactions data, since coal prices are available only for deliveries at the plant. Mine-mouth coal prices and transportation charges are confidential. Our hedonic regressions indicate negative implicit prices of sulfur and positive implicit prices of Btu. To our knowledge, we provide the first estimates of implicit prices for sulfur and Btu in coal net of confidential transportation charges.[7] Cost

[6] One exception is Atkinson and Tsionas (2016).

[7] Atkinson and Dorfman (2005) and Färe, Grosskopf, Noh, and Weber (2005), among others, have computed relative shadow prices per ton of SO_2 (which imply estimates of implicit prices of sulfur). However, they estimate relative implicit prices independent of the price of coal using duality theory, rather than a hedonic regression which nets out transportation charges from the delivered price of fuel.

function estimates find increasing marginal abatement costs, which we validate using an engineering simulation model. The estimated production function parameters indicate moderate increasing returns to scale, which is consistent with the literature as reviewed in Atkinson (2018). We then compare our estimates with those obtained using the LP/ACF methodology, finding that the traditional approach and our cost function approach produce substantially different parameter estimates.

A complete data set for our plants ends in 2005, due to subsequent relaxed reporting requirements. Since a key goal of the Acid Rain Program was to encourage production efficiency[8], we use our estimates of plant production and abatement efficiencies together with the sulfur and Btu content of coal, among other variables, to explain plant closings from 2006-2017. During this period, our sample plants retired 45% of their generating units. Weak evidence exists that greater production efficiency increases the probability of survival for units with flue gas desulfurization (FGD) units and non-FGD units. However, for the latter units we find that greater relative abatement efficiency significantly increases the probability of plant survival. We also find that higher Btu content, and more importantly, higher sulfur content significantly lessen the probability of survival.

We examine two counterfactual cases motivated by recently proposed and implemented market interventions by the Administration designed to bring back coal. The first counterfactual considers the effect of a coal-cost subsidy by compensating plants for maintaining a 90-day supply of coal on hand. We find that this compensation, equivalent to a 25% subsidy of the total cost of coal, causes FGD and non-FGD plants to reduce the sulfur content by 15.97% and 28.22%, respectively. They would also reduce the Btu content by 2.08% and 2.72%, respectively. Further, SO_2 emissions decrease by approximately 26% for both types of plants. The second counterfactual models the current Administration's plan

[8] See the Environmental Protection Agency web site:https://www.epa.gov/airmarkets/acid-rain-program.

to dramatically slow the reduction in SO_2 emissions relative to levels expected under the Clean Power Plan. We considers the effect of a 20% increase in total SO_2 emission allowances, within the context of the current non-functional SO_2 permit trading market. We find that FGD plants would reduce sulfur content by 11.43%, while non-FGD plants would increase it by 13.01%. Their Btu content would decrease by 2.23% and 2.72%, respectively. The FGD plants would save 67.38% on abatement costs and 2.22% on coal costs, while non-FGD plants would save 2.59% on coal costs. Ironically, the first measure substantially reduces sulfur content and SO_2 emissions from both FGD and non-FGD units, while the second measure reduces average sulfur content across plant types.

2 Data

Our data is a balanced panel of the 80 largest coal-fired power plants in the U.S. from 1995 to 2005.[9] Our sample ends in 2005 since an increasing number of private utilities did not report capital and labor data after 2005. The majority of our sample plants are located in the Southern, Mid-Atlantic, or Midwestern states, with a few in the Rocky Mountain and Far Western regions. The technology modeled in this study consists of the inputs capital, labor, Btu and sulfur from coal, which produce megawatt hours (mWh) and the pollutant sulfur dioxide (SO_2). Capital is measured as megawatt (MW) generating capacity of the plant, which is adjusted by the plant as existing capacity is augmented with new capital or reduced through depreciation of old capital. Labor is the number of full-time employees plus one-half the number of part-time employees. The sulfur content of the coal burned is measured as a percentage. While the power plants in our sample consume coal and either oil or natural gas, on average 99% of the Btu generated by each plant comes from coal consumption.

[9] See the Appendix for the list of plants.

We obtained our data from a number of different sources. FERC Form 1 provides labor and capital data for private electric power plants, and the EIA-412 survey is the source of this data for public power plants. While DOE halted the EIA-412 survey after 2003, the Tennessee Valley Authority voluntarily posted 2004-06 data for its electric power plants on-line. The U.S. Department of Energy (DOE) Form EIA-767 survey is the source of information about fuel consumption and net mWh generation by plant. The SO_2 emissions data at the plant level are collected by the EPA as part of its Continuous Emissions Monitoring System. The sulfur content and Btu content of the coal burned by plant comes from EIA-423.[10]

We have data on the price of coal as delivered to each plant. From EIA Form 423 we obtain the price of deliveries of coal to the plant. Although we have prices of capital and labor inputs only at the utility level, we make the reasonable assumption that plant-level prices are identical to firm-level prices for these inputs, since these prices are determined at the utility level. We compute the user cost of capital at the firm level using the corporate tax rate, the corporate property tax rate, the depreciation rate, the firm's yield, and the Handy-Whitman Index as in Atkinson, Primont, and Tsionas (2015). The yield on the firm's latest issue of long-term debt comes from Moody's Public Utility Manual (before 2001) and from Mergent's Public Utility Manual after that time. From FERC Form 1 we collect the wage paid by the firm as salaries plus wages for electric operating and maintenance workers divided by the number of full time workers plus one-half the number of part-time workers for the firm.

We also collected a number of variables that measure coal quality and environmental costs. These include the SO_2 removal rate of scrubbers, the percent of total plant MW capacity that is scrubbed, and finally the O&M and capital costs of FDG devices. The

[10]We wish to thank Carl Pasurka for supplying us with data on input and output quantities.

O&M costs of the FGD operation include the cost of the feed materials and chemicals, FGD labor, waste disposal, and other costs. These data come from EIA Forms 767 and 860.

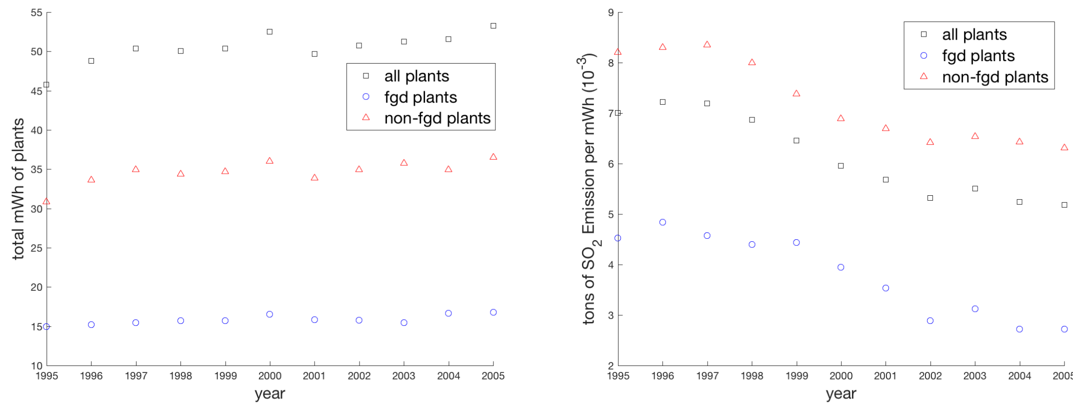
Table 1 shows the summary statistics of the data. Among the 80 plants, only 18 plants employ FGD units in all years from 1995 to 2005. The other 62 plants have either never installed FGD units or installed them for only part of this period. The left panel is for non-FGD plants and the right panel is for FGD plants. The FGD plants are larger than the non-FGD ones on average, with greater electrical generation, capital, labor, and coal-consumption. They use coal with a significantly higher sulfur content and slightly lower Btu content than non-FGD plants. The median sulfur content is 1.478% for FGD plants, but only 0.887% for non-FGD plants. The median Btu content is 22.521 million/ton for FGD plants and 24.144 million/ton for non-FGD plants. The median coal price is lower for FGD plants because of higher sulfur and lower Btu. Median coal prices are $25.619/ton for FGD plants and $38.065/ton for non-FGD plants. The median sulfur removal rate of scrubbed capacity is 85% for FGD plants with large variation across plants. The capital yield and the labor wage are similar for the two types of plants. The median capital yields are 7.55% for non-FGD plants and 7.54% for FGD plants, while median wages are $43,560 for non-FGD plants and $43,339 for FGD plants.

Table 1: Data Summary Statistics for Plants

	No FGD			FGD		
variable	median	min	max	median	min	max
Generation (10^6mwh)	4.417	0.116	22.329	7.663	1.860	20.321
Generation Capacity (MW)	772	110	3,498	1,620	411	2,600
Labor	129	23.744	578	212	64.634	538
Coal (10^6 tons)	1.967	0.610	12.308	3.999	0.865	9.135
Sulfur (%)	0.887	0.125	3.788	1.478	0.326	3.947
Btu (10^6/ton)	24.144	16.212	26.348	22.521	15.451	24.639
Removal (%)				85.041	37.021	97.700
SO_2 emission (tons)	28,209	631	186,470	21,651	3,242	212,377
Yield (%)	7.550	5.380	8.950	7.54	5.380	97.700
Wage (10^4\$)	4.356	2.491	9.468	4.339	2.675	8.320
Coal price (\$/ton)	38.065	11.382	141.481	25.619	9.473	53.348
FGD O&M costs (10^3\$)				3,793	300	30,015
N	682	682	682	198	198	198

Generation capital has been slowly increasing for all plants. The average annual growth of plant capital is 3.66% for all plants. The level of labor used in generation has been decreasing for all plants, with an annual growth of −4.06%. Plant-level heat input from coal has increased slightly over time, with a mean growth rate of 0.96%. Plant-level electricity generation growth rate is close to the heat growth rate, with a mean of 0.96%. Over our sample period, total electricity generated has increased, while the SO_2 emitted per mWh of electricity has fallen for both non-FGD and FGD plants. Average electricity generation in 10^6 mWh increased from 5.711 in 1995 to 6.647 in 2005.

In Figure 1a, we represent the total generation for all plants with squares, for non-FGD plants with triangles, and for FGD plants with circles. Non-FGD plants generate about twice as much electricity as their FGD counterparts due to large number of non-FGD plants. In Figure 1b, we represent the average SO_2 per mWh generation for all plants with squares, for FGD plants with circles, and for non-FGD plants with triangles. We see that SO_2 production in thousands of short tons per mWh has fallen substantially for both types of plants. The decline is almost 50% for FGD plants and 25% for non-FGD plants. The non-FGD plants emit about 40-80% more SO_2 per mWh of electricity than the FGD plants.

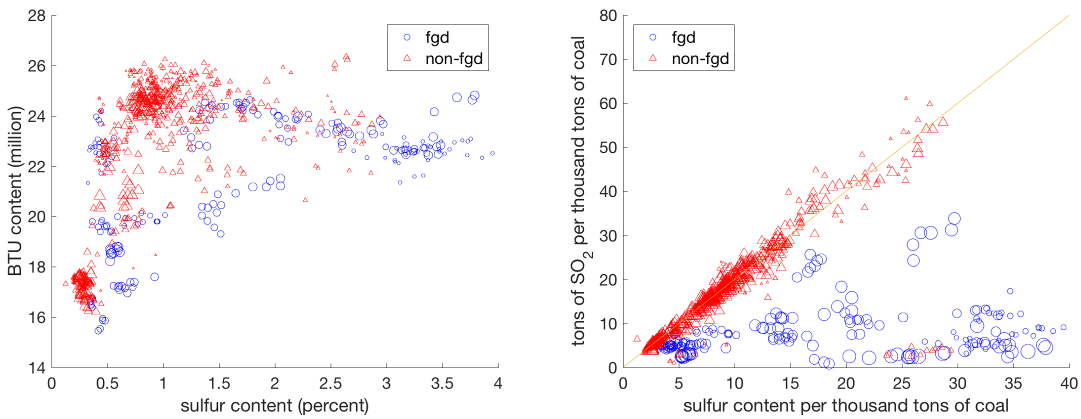


(a) Total generation by plant type　　(b) SO_2 Emission per mWh Generation

Figure 1: Total Generation and SO_2 Emission per mWh Generation

In Figure 2a, we plot the plant-year Btu content against sulfur content, indicating a wide variety of combinations of them in the upper-triangular portion. We represent the non-FGD plants with triangles and the FGD plants with circles, where the size of each indicates the magnitude of coal purchases. Plants possess considerable flexibility in trading off these two characteristics of coal. The range of substitution possibilities is greater for FGD plants since they have the option of mixing fuel types and employing FGD. The lower triangular part of this graph is either avoided by plants or not available for purchase.

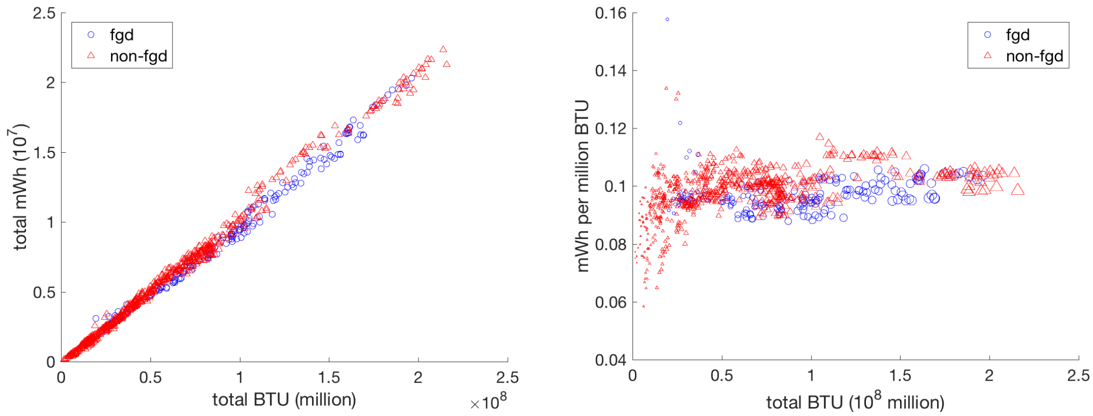
In Figure 2b, we plot the plant-year generation of SO_2 against sulfur content. As expected, for the non-FGD plants, there is a linear relationship between the two variables with a slope of approximately 2.0, as expected from the chemistry of converting sulfur into SO_2 without controls. The FGD plants exhibit a wide range of differences in the percent of emissions per ton relative to sulfur content, since plants differ substantially in the percent of emissions that are controlled as well as the control efficiencies of FGD devices as seen from Table 1. This figure also indicates the considerable substitution possibilities among Eastern coal mines and more limited but still substantial substitution possibilities among Western mines. The former produce higher Btu/higher sulfur coal, while the latter produce lower Btu/lower sulfur coal. We also observe in the data that the average sulfur content and Btu content of all 80 coal-fired power plants have decreased from 1995 to 2005.



(a) Btu Content and Sulfur Content

(b) SO_2 Emission and Sulfur Content

Figure 2: Sulfur Content, Btu Content, and SO_2 Emission



(a) kWh generation versus Btu consumption (b) Heterogeneity in Generation per million Btu

Figure 3: kWh Generation versus Btu Consumption

In Figure 3a, we graph total generation against total Btu for non-FGD and FGD plants. Figure 3a shows a close to linear relationship between generation and heat input. Nonetheless, there exists great variation in electricity generation per million Btu across plants and years, as shown in Figure 3b.

3 Cost Minimization Problem of Coal Plants

We model the choice of Btu and sulfur content for a set of coal-fired electricity generation plants, which are attempting to minimize the costs of coal consumption and pollution control subject to constraints on total generation and environmental degradation.

3.1 Production Costs

In period t, plant j first observes its generation capacity available to produce electricity, k_{jt}, its labor stock, l_{jt}, its generation productivity, ω^y_{jt}, and its exogenously determined target output, y_{jt}, and then chooses the quality and the quantity of coal to produce y_{jt}. The two key quality characteristics of coal are its Btu content per ton of coal, b_{jt}, and its sulfur content per ton of coal, s_{jt}. Both affect the total cost of coal and the cost of pollution control for a plant. First, both b_{jt} and s_{jt} affect the coal price, which theoretically increases with b_{jt} and decreases with s_{jt}. Second, given the output level, the amount of coal a plant needs depends b_{jt}. The higher b_{jt}, the less coal the plant needs to consume. Lastly, b_{jt} and s_{jt} both affect the abatement cost to control SO_2. While b_{jt} determines the total amount of coal for a plant, s_{jt} determines the amount of sulfur per ton of coal.

Let the total Btu consumed (total heat input) be h_{jt} and assume that the plant's non-stochastic production function for electricity has a Cobb-Douglas form,

$$y_{jt} = e^{(\beta_0 + \omega^y_{jt})} k_{jt}^{\beta_k} l_{jt}^{\beta_l} h_{jt}^{\beta_h}, \tag{1}$$

where $(\beta_0, \beta_l, \beta_k, \beta_h)$ are parameters. ω^y_{jt} measures the heterogenous unobserved production efficiency, which can be from the input quality difference if such difference exists or the overall operating efficiency difference of the plants. For example, if a plant has more productive employees or can generate more electricity with the same input bundle than average, then it has a higher ω^y_{jt}. Given the production function in (1), the h_{jt} needed to produce y_{jt} for a given $(k_{jt}, l_{jt}, \omega^y_{jt})$ is

$$h_{jt}(y_{jt}, l_{jt}, k_{jt}, \omega^y_{jt}) = (y_{jt} e^{-(\beta_0 + \omega^y_{jt})} l_{jt}^{-\beta_l} k_{jt}^{-\beta_k})^{\frac{1}{\beta_h}}. \tag{2}$$

Total h_{jt} decreases as the productivity ω^y_{jt} increases or as the output y_{jt} decreases. If we apply the traditional approach of OP/LP/ACF, inverting this function would yield ω^y_{jt} as a function of $(k_{jt}, l_{jt}, h_{jt}, y_{jt})$. Substituting this into (1) would yield a non-viable estimation equation, since output would be a function of output. Instead, we formulate our cost function approach.

For any b_{jt}, the tons of coal required to produce y_{jt} by plant j is

$$n(b_{jt}; y_{jt}, l_{jt}, k_{jt}, \omega^y_{jt}) = \frac{h_{jt}(y_{jt}, l_{jt}, k_{jt}, \omega^y_{jt})}{b_{jt}} = (y_{jt} e^{-(\beta_0 + \omega^y_{jt})} l_{jt}^{-\beta_l} k_{jt}^{-\beta_k})^{\frac{1}{\beta_h}} b_{jt}^{-1}. \qquad (3)$$

The higher the b_{jt}, the less amount of coal the plant must consume to produce the given electricity output. The price of coal per ton as delivered to the utility plant, w^c_{jt}, is a function of b_{jt}, s_{jt}, and freight or shipping costs per ton, f_{jt}, from the mine to the plant. Thus, $w^c_{jt} = w^c_{jt}(b_{jt}, s_{jt}, f_{jt})$. The total cost of coal is

$$w^c_{jt}(b_{jt}, s_{jt}, f_{jt}) n(b_{jt}; y_{jt}, l_{jt}, k_{jt}, \omega^y_{jt}).$$

3.2 Pollution Control Costs

The total pollution control cost function for FGD plants can be very different from that for non-FGD plants. For an FGD plant, total pollution control cost includes the expenditures on FGD to control emissions and on the purchase of SO_2 pollution permits (and the opportunity cost of any allowances held by the plant) for uncontrolled emissions. All uncontrolled emissions require an emission permit. A non-FGD plant solely relies on emission permits to comply with environment regulation. Its pollution control cost includes only the expenditures to purchase permits and the opportunity cost of any allowances held by the plant.

3.2.1 Pollution Control Costs of Plants with Scrubbers

The pollution control cost of an FGD plant includes the SO_2 abatement cost and the cost of emission. The total abatement cost depends on the amount of coal used in generation units that have scrubbers, $n^a_{jt}(\leq n_{jt})$, the sulfur content, s_{jt}, the sulfur removal efficiency, r_{jt}, and its abating efficiency, ω^a_{jt}. Denote the total abatement cost function by $C^a(n^a_{jt}, s_{jt}, r_{jt}, \omega^a_{jt})$.

The abatement cost is expected to increase with the amount of sulfur scrubbed, $n^a_{jt}s_{jt}r_{jt}$. It will be nonlinear if the marginal pollution control cost changes with the arguments. We assume that the abatement cost is a power function of the total amount of sulfur scrubbed:

$$C^a(n^a_{jt}, s_{jt}, r_{jt}, \omega^a_{jt}) = e^{\lambda_0 - \omega^a_{jt}} F(n^a_{jt}, s_{jt}, r_{jt}) = e^{\lambda_0 - \omega^a_{jt}} (n^a_{jt} s_{jt} r_{jt})^{\lambda}. \tag{4}$$

The constant λ_0 measures the average log abatement cost that is the same for all plants. Because the abatement cost includes the cost of the feed materials and chemicals, waste disposal, FGD labor, and other costs, ω^a_{jt} represents the heterogeneity in these costs across plants. For example, if a plant faces higher costs, then the plant has a lower abating efficiency, ω^a_{jt}.

This plant needs to either use its free permit allowances or buy permits to emit SO_2 without abating. Given the amount of sulfur abated is $n^a_{jt}s_{jt}r_{jt}$, the remaining sulfur in the coal is $s_{jt}(n_{jt} - n^a_{jt}r_{jt})$ (tons), which will be turned into $S^e_{jt} = 2s_{jt}(n_{jt} - n^a_{jt}r_{jt})$ tons of SO_2 (based on molecular weight) assuming that all sulfur is transformed into SO_2. Figure 2b shows that the total weight of SO_2 is about two times the weight of total sulfur for non-FGD plants. Since a plant can trade its endowed free allowances, there is also an opportunity cost to holding allowances. Let the permit price faced by plant j in year t be p_{jt}. Then the cost of buying permits for the emission for plant j in year t is

$$p_{jt}S^e_{jt} = 2p_{jt}s_{jt}(n_{jt} - n^a_{jt}r_{jt}). \tag{5}$$

Therefore, total pollution control cost is the sum of the costs of using scrubbers and the cost of purchased permits/allowances.

3.2.2 Tradeoff between Abating and Emitting Sulfur Dioxide

A plant with FGD units incurs total pollution control costs equal to the sum of scrubbing costs and the costs of buying permits. Since this plant optimally chooses how much coal to use in units with scrubbers, the minimization problem for total pollution control costs is

$$\min_{n_{jt}^a} \left\{ e^{\lambda_0 - \omega_{jt}^a} (n_{jt}^a s_{jt} r_{jt})^\lambda + 2p_{jt} s_{jt} (n_{jt} - n_{jt}^a r_{jt}) \right\}. \tag{6}$$

If $\lambda > 1$, then the marginal cost of abatement increases with the scrubbed coal quantity, which implies a unique optimal abatement level is determined by equating the marginal abatement cost to the marginal emission cost. The optimal amount of coal that should be scrubbed, n_{jt}^{a*}, must satisfy the first-order condition (FOC):

$$e^{(\lambda_0 - \omega_{jt}^a)} \lambda (n_{jt}^a)^{\lambda - 1} (s_{jt} r_{jt})^\lambda - 2p_{jt} s_{jt} r_{jt} = 0.$$

Thus, the optimal amount of coal to abate for plant j in year t is a function of $(s_{jt}, r_{jt}, p_{jt}, \lambda, \omega_{jt}^a)$:

$$n_{jt}^{a*} = \Big(\frac{2p_{jt} e^{\omega_{jt}^a - \lambda_0}}{\lambda} \Big)^{\frac{1}{\lambda - 1}} \frac{1}{s_{jt} r_{jt}}. \tag{7}$$

The optimal abatement amount increases with ω_{jt}^a and p_{jt} and decreases with r_{jt}, s_{jt}, and λ. Plugging n_{jt}^{a*} into the pollution control cost function, we obtain the minimum pollution control cost as a function of n_{jt} and s_{jt}, among other variables. Because n_{jt} depends on y_{jt}, b_{jt}, and other plant-year specific variables as in equation (3), we can write the pollution control cost function in terms of $b_{jt}, s_{jt}, \omega_{jt}^a, \omega_{jt}^y$, and $X_{jt} = (r_{jt}, p_{jt}, y_{jt}, l_{jt}, k_{jt})$:

$$C_{\mathrm{FGD}}^s(s_{jt}, b_{jt}; X_{jt}, \omega_{jt}^y, \omega_{jt}^a) = \Big(\frac{1}{\lambda} - 1 \Big) 2p_{jt} \Big(\frac{2p_{jt} e^{\omega_{jt}^a - \lambda_0}}{\lambda} \Big)^{\frac{1}{\lambda - 1}} + 2p_{jt} s_{jt} \frac{h_{jt}(y_{jt}, l_{jt}, k_{jt}, \omega_{jt}^y)}{b_{jt}}. \tag{8}$$

The last term is the cost that plants with FGD units would incur if they performed no control and used permit allocations and purchased permits to cover all emissions. However, they can reduce their total control costs below this level. This cost savings is given by the first term in (8), which is equal to $e^{\lambda_0-\omega^a_{jt}}(n^a_{jt}s_{jt}r_{jt})^\lambda - 2p_{jt}s_{jt}n^a_{jt}r_{jt}$ from equation (6). This is the difference between the cost of control for the abated SO_2 minus the expenditure that would have been incurred on SO_2 permits and allowances if these emissions were unabated. If the marginal cost of control is increasing, $\lambda > 1$, then this difference is negative. That is, abatement expenditures by the FGD plant are less than what they would have been if the plant had relied solely on permit allocations or purchased permits. Hence, the FGD plant spends less on pollution control than the non-FGD plant, which can only use permit allocations or buy permits to cover emissions.

3.2.3 Pollution Control Cost of Plants without Scrubbers

For a plant without scrubbers, the amount of SO_2 generated and emitted is $S^e_{jt} = 2s_{jt}n_{jt}$ tons. The cost of permits to emit S^e_{jt} is the plant-year specific permit price times the amount of emission.

$$C^s_{\text{NFGD}}(s_{jt}, b_{jt}; X_{jt}, \omega^y_{jt}) = p_{jt}S^e_{jt} = 2p_{jt}s_{jt}n_{jt} = 2p_{jt}s_{jt}\frac{h_{jt}(y_{jt}, l_{jt}, k_{jt}, \omega^y_{jt})}{b_{jt}}, \tag{9}$$

which increases with the sulfur content and decreases with the Btu content.

3.3 Minimization of Total Variable Cost

3.3.1 Plants with Scrubbers

A plant minimizes the sum of its total production cost and abatement cost by choosing b_{jt} and s_{jt}. The plant faces constraints on (b_{jt}, s_{jt}) for several reasons. First, if the plant has signed long-term contracts over time with coal mines, then the plant can only partially adjust the coal characteristics in a given year. Second, the range of (b_{jt}, s_{jt}) that a plant faces depends on the availability of coal.[11]

$$\min_{b_{jt}, s_{jt}} w^c_{jt}(b_{jt}, s_{jt}; f_{jt}) n(b_{jt}; X_{jt}, \omega^y_{jt}) + C^s_{\text{FGD}}(s_{jt}, b_{jt}; X_{jt}, \omega^y_{jt}, \omega^a_{jt}),$$

subject to the plant-year specific constraints on (b_{jt}, s_{jt})

$$b_{jt} \in \left[\underline{b}_{jt}, \bar{b}_{jt}\right],$$
$$s_{jt} \in \left[\underline{s}_{jt}, \bar{s}_{jt}\right],$$

where the underbars and overbars indicate lower and upper bounds.

We model the optimal choice of (b_{jt}, s_{jt}), conditional on the observed shipping cost f_{jt}. The assumption is that, for the given f_{jt}, each plant can choose from a continuum of (b, s). Thus, we can model the choice of (b, s) as a continuous one. This assumption is reasonable because a plant can purchase coal with quite different Btu and sulfur content either at a given mine or at adjacent mines.[12] Therefore, a plant can choose different combinations of (b, s) for the same shipping cost.[13]

[11] Figure 2a shows that coal with very high-sulfur and very low-Btu content has not been purchased in the US.

[12] Appendix C of DOE (1999) for coal reserves and EIA Forms 423 and 923 (1995-2015) for coal deliveries provide evidence that a wide variety of coal with varying Btu and sulfur content is both available in reserves and has been sold by mines in each major coal-producing state. For example, Pennsylvania and West Virginia have proven coal reserves with sulfur content ranging from .5 to 3.0 percent in combination with Btu ranging from 23 to 26 MMBtu per ton. Mines in the Powder River Basin have proven reserves with sulfur content ranging from .1 to 2.5 percent in combination with Btu ranging from 15 to 23 MMBtu per ton.

[13] We do not model the choice of f_{jt} because the shipping cost per ton of coal and the distance for each coal shipment are confidential.

The FOC with respect to b_{jt} is

$$\frac{\partial w_{jt}^c(b_{jt}, s_{jt}; f_{jt})}{\partial b_{jt}} n_{jt} + w_{jt}^c(b_{jt}, s_{jt}; f_{jt}) \frac{\partial n_{jt}}{\partial b_{jt}} + \frac{\partial C_{\text{FGD}}^s(s_{jt}, b_{jt}; X_{jt}, \omega_{jt}^y, \omega_{jt}^a)}{\partial b_{jt}} - \underline{\mu}_{jt}^b + \bar{\mu}_{jt}^b = 0, \tag{10}$$

where $(\underline{\mu}_{jt}^b, \bar{\mu}_{jt}^b)$ are the plant-year specific Lagrangian multipliers for lower- and upper-bound constraints for Btu content. The first term measures the marginal cost of the Btu content. The second term is the marginal savings in the cost of coal because less coal is consumed when choosing higher Btu-content coal. For the same reason, a higher Btu content also reduces the pollution control cost for a given s_{jt}. The third term measures the marginal savings in the pollution control cost. The first-order condition with respect to s_{jt} is

$$\frac{\partial w_{jt}^c(b_{jt}, s_{jt}; f_{jt})}{\partial s_{jt}} n_{jt} + \frac{\partial C_{\text{FGD}}^s(s_{jt}, b_{jt}; X_{jt}, \omega_{jt}^y, \omega_{jt}^a)}{\partial s_{jt}} - \underline{\mu}_{jt}^s + \bar{\mu}_{jt}^s = 0, \tag{11}$$

where $(\underline{\mu}_{jt}^s, \bar{\mu}_{jt}^s)$ are the plant-year specific Lagrangian multipliers for lower- and upper-bound constraints for sulfur content. The first term is the marginal reduction in the price of coal from choosing a higher sulfur content, while the second term is the marginal abatement cost of using higher sulfur content coal.

Using the pollution control cost function for FGD plants, we can derive the explicit functional forms of the two FOCs. We have

$$\frac{\partial w_{jt}^c}{\partial b_{jt}} - \frac{w_{jt}^c}{b_{jt}} - \frac{2 p_{jt} s_{jt}}{b_{jt}} - \frac{\underline{\mu}_{jt}^b}{n_{jt}} + \frac{\bar{\mu}_{jt}^b}{n_{jt}} = 0. \tag{12}$$

$$\frac{\partial w_{jt}^c}{\partial s_{jt}} + 2 p_{jt} - \frac{\underline{\mu}_{jt}^s}{n_{jt}} + \frac{\bar{\mu}_{jt}^s}{n_{jt}} = 0. \tag{13}$$

Denote the optimal b_{jt} and s_{jt} by (b_{jt}^*, s_{jt}^*). We can write the total variable cost function as

$$C(X_{jt},\omega^y_{jt},\omega^a_{jt},p_{jt},\lambda)=\big(w^c_{jt}(b^*_{jt},s^*_{jt};f_{jt})+2p_{jt}s^*_{jt}\big)\frac{h_{jt}(X_{jt},\omega^y_{jt})}{b^*_{jt}}+\Big(\frac{1}{\lambda}-1\Big)2p_{jt}\Big(\frac{2p_{jt}e^{\omega^a_{jt}}}{\lambda}\Big)^{\frac{1}{\lambda-1}}. \tag{14}$$

3.3.2 Plants without Scrubbers

The non-FGD plants choose (b_{jt}, s_{jt}) to minimize the total variable cost of coal and permits.

$$\min_{b_{jt},s_{jt}} w^c_{jt}(b_{jt},s_{jt};f_{jt})n(b_{jt};X_{jt},\omega^y_{jt})+C^s_{\mathrm{NFGD}}(s_{jt},b_{jt};X_{jt},\omega^y_{jt}),$$

subject to the constraints on (b_{jt}, s_{jt})

$$\begin{aligned} b_{jt} &\in \left[\underline{b}_{jt},\bar{b}_{jt}\right], \\ s_{jt} &\in \left[\underline{s}_{jt},\bar{s}_{jt}\right]. \end{aligned}$$

Using the permit cost in equation (9), the FOCs for b_{jt} and s_{jt} are

$$\begin{aligned} \frac{\partial w^c_{jt}}{\partial b_{jt}}-\frac{w^c_{jt}}{b_{jt}}-\frac{2p_{jt}s_{jt}}{b_{jt}}-\frac{\underline{\mu}^b_{jt}}{n_{jt}}+\frac{\bar{\mu}^b_{jt}}{n_{jt}} &= 0. \\ \frac{\partial w^c_{jt}}{\partial s_{jt}}+2p_{jt}-\frac{\underline{\mu}^s_{jt}}{n_{jt}}+\frac{\bar{\mu}^s_{jt}}{n_{jt}} &= 0. \end{aligned}$$

The non-FGD plants have the same FOCs as the FGD plants. Denote the optimal b_{jt} and s_{jt} by (b^*_{jt}, s^*_{jt}), the total variable cost for the plant is

$$C(X_{jt},\omega^y_{jt})=(w^c_{jt}(b^*_{jt},s^*_{jt};f_{jt})+2p_{jt}s^*_{jt})\frac{h_{jt}(X_{jt}\omega^y_{jt})}{b^*_{jt}}. \tag{15}$$

4 Econometric Model and Estimation

Our estimation of the model parameters consists of three steps. First, we estimate the coal price as a function of the Btu content and sulfur content of coal using the coal transaction level data for all the plants in our sample. Using the estimated price function, we compute the marginal price of Btu content and sulfur content for each plant-year observation. Second, we estimate the abatement cost function using plant-year level data on the FGD O&M cost and the sulfur abatement level. Lastly, we estimate the production function parameters using the total variable cost function with plant-year level data. In the second and the third steps, we add the transition functions of the unobserved efficiencies to help identify the parameters. This follows the traditional literature on production function literature.

4.1 Estimation of Coal Price Function

The delivered price of a ton of coal depends on the Btu content, the sulfur content, the plant fixed effect, the mine fixed effect, the year fixed effect, the distance between the mine, the amount of coal purchased, the total amount of allowances for the whole country in a year, and the contract type of the purchase.[14] To control for the freight charges on transporting the coal, we use the mine dummy, the plant dummy, and the interaction of them for each coal shipment.

We assume that the plants face coal price functions that are plant-year specific and vary with source mines, $w^c_{jt}(b, s; f,) = \overline{w}^c_{jt}(b, s) + w_{jt}(f)$. The stochastic hedonic coal price

[14] There are three types of coal purchase contracts depending on the length of the contract and whether the contract is new, C, NC, and S. Type C contracts are for purchases received under a purchase order or contracts that has a duration of one year or longer. Type NC contracts are new contracts or renegotiated contract purchases under which deliveries were first made during the reporting month. Type S are for the spot-market purchases of coal received under a purchase order or contract that has a duration of less than one year.

function is

$$\begin{aligned} w^c_{jt}(b_{jt}, s_{jt}; f_{jt}) = {} & \alpha_0 + \alpha_1 b_{jt} + \alpha_2 s_{jt} + \alpha_3 b^2_{jt} + \alpha_4 s^2_{jt} + \alpha_5 b_{jt} * s_{jt} \\ & + \alpha_6 log(n_{jt}) + \alpha_7 s_{jt} * A_t + \sum_{q=1}^{Q} \alpha_{8q} d_q + \sum_{t=1}^{T} \alpha_{9t} d_t \\ & + \sum_{m=1}^{M} \alpha_{10,m} d_m + \sum_{k=1}^{J} \alpha_{11,k} d_k + \sum_{m=1}^{M} \sum_{k=1}^{J} \alpha_{12,mk} d_m d_k + \epsilon^w_{jt}, \end{aligned} \tag{16}$$

where d_q is the dummy for the contract type, d_t is the year dummy for year $t, t = 1, \ldots, T$, d_m is the mine dummy $m, m = 1, \ldots, M$, that sells coal to plant j, and d_k is the plant dummy. A_t is the total amount of SO_2 allowances in the United States in year t.[15] ϵ^w_{jt} is a coal price shock. The term $log(n_{jt})$ represents possible quantity discount.

The stochastic coal price equation (16) can be estimated using coal transaction data directly via OLS regression. Each observation is a transaction between a mine and a plant. There could be multiple transactions between a mine and a plant in a year. After estimating the coal price function, we aggregate the transaction level data to get the plant-year level average b_{jt} and s_{jt}, weighted by the coal quantity. We evaluate the plant-year level marginal prices of Btu content and sulfur content as

$$\frac{\partial w^c_{jt}(b_{jt}, s_{jt}; f_{jt})}{\partial b_{jt}} = \alpha_1 + 2\alpha_3 b_{jt} + \alpha_5 s_{jt}, \tag{17}$$

$$\frac{\partial w^c_{jt}(b_{jt}, s_{jt}; f_{jt})}{\partial s_{jt}} = \alpha_2 + 2\alpha_4 s_{jt} + \alpha_5 b_{jt} + \alpha_7 A_t. \tag{18}$$

The marginal prices of coal will be used in the FOCs of the cost minimization problem.

After estimating the coal price function, we compute the plant-year specific permit prices using the FOCs for the optimal s_{jt} for both types of plants in equations (12) and

[15] The total amount includes both the new allowances issued in that year and the allowances banked from previous years.

(13). More specifically, we argue that $\underline{\mu}^s_{jt} = \bar{\mu}^s_{jt} = 0$, $\underline{b}_{jt}$ and $\bar{b}_{jt}$ can not both be zero. From the coal transaction level data, we know that every plant bought coal from multiple mines in a year, so every plant has some flexibility in choosing (b_{jt}, s_{jt}), which implies that $\underline{b}_{jt} < \bar{b}_{jt}$ and $\underline{s}_{jt} < \bar{s}_{jt}$. Thus, at least one constraint for the Btu(sulfur) content is not binding, so at least one $\mu^b_{jt} \in \{\underline{\mu}^b_{jt}, \bar{\mu}^b_{jt}\}$ and at least one $\mu^s_{jt} \in \{\underline{\mu}^s_{jt}, \bar{\mu}^s_{jt}\}$ are zero.

There are four cases for the values of the Lagrangian multipliers that satisfy this condition. First, none of the four constraints is binding. That is, $\underline{\mu}^s_{jt} = \bar{\mu}^s_{jt} = \underline{\mu}^b_{jt} = \bar{\mu}^b_{jt} = 0$. Plugging them into the FOCs, we find that the two FOCs imply very different values for p_{jt}, which implies that at least one constraint is binding. Second, the two constraints for Btu content are not binding, and one constraint for sulfur content is binding. That is, $\underline{\mu}^b_{jt} = \bar{\mu}^b_{jt} = 0$, and either $\underline{\mu}^s_{jt} = 0$ or $\bar{\mu}^s_{jt} = 0$. Plugging them into the FOCs, we can get p_{jt}, but these permit prices are dramatically different from the yearly average permit prices in the permit auction data. Third, the two constraints for sulfur content are not binding, and one constraint for Btu content is binding. That is, $\underline{\mu}^s_{jt} = \bar{\mu}^s_{jt} = 0$, and either $\underline{\mu}^b_{jt} = 0$ or $\bar{\mu}^b_{jt} = 0$. Plugging them into the FOCs, we get permit prices that are reasonably close to the auction prices.[16] Lastly, one constraint for Btu content is binding, and one constraint for sulfur content is binding. This implies that the plant chooses corner solutions for both Btu content and sulfur content, but this is unlikely given the fact that the plants are buying coal from multiple mines with a variety of coal characteristics each year. Therefore, the third case is the most consistent to the data, and $p_{jt} = -2\frac{\partial w^c_{jt}(b_{jt}, s_{jt}; f_{jt})}{\partial s_{jt}}$ in this case. This finding also alignes with the variation of Btu content and sulfur content in the data. The standard deviation of Btu content across plants is only 1.93% of the average Btu content, but the standard deviation of sulfur content across plants is 13.13% of the average sulfur content. Thus, plants are more likely to face binding constraints on Btu content than sulfur

[16] See Appendix 8 for the comparison of the permit prices.

content.

4.2 Estimation of the Abating Cost Function

To capture the potential consistency in the plants' abating efficiency, we assume that a plant's current abating efficiency affects its abating efficiency in the next year.

$$\omega^a_{jt+1} = g^a(\omega^a_{jt}) + \xi^a_{jt+1} = \rho^a_0 + \rho^a_1\omega^a_{jt} + \rho^a_2(\omega^a_{jt})^2 + \xi^a_{jt+1}, \tag{19}$$

where ξ^a_{jt+1} is the shock to the abating efficiency. Specifically, ξ^a_{jt+1} represents the shock to the plant-year specific costs of FGD feed, waste disposal, labor, and other costs.

The parameters to be estimated in this step are λ in equation (4) and the $\rho^a = (\rho^a_0, \rho^a_1, \rho^a_2)$ in equation (19). The logarithm of abatement cost is

$$\log C^a(s_{jt}, n^a_{jt}, \omega^a_{jt}) = \lambda_0 - \omega^a_{jt} + \lambda(\log s_{jt} + \log n^a_{jt} + \log r_{jt}) + \epsilon^a_{jt},$$

where ϵ^a_{jt} is an idiosyncratic error of the abatement cost. We assume ϵ^a_{jt} is mean zero and uncorrelated to the plant's coal prices and capacity.

Because the abating efficiency ω^a_{jt} is unobserved, we use its transition function and the optimal quantity of coal abatement to evaluate it. Replacing the abating efficiency in the log abating cost function by $\omega^a_{jt} = g^a(\omega^a_{jt-1}) + \xi^a_{jt}$, we get

$$\log C^a(s_{jt}, n^{a*}_{jt}, \omega^a_{jt}) = \lambda_0 - g^a(\omega^a_{jt-1}) + \lambda(\log s_{jt} + \log n^a_{jt} + \log r_{jt}) - \xi^a_{jt} + \epsilon^a_{jt}. \tag{20}$$

From the optimal quantity of abated coal in (7), we know that $\omega^a_{jt-1} = (\lambda - 1)\log(s_{jt-1}n^a_{jt-1}r_{jt-1}) + \log\left(\frac{\lambda}{2p_{jt-1}}\right)$. Plugging this into equation (20), we can replace the unobserved abating efficiency with observed variables. The new error term is $(-\xi^a_{jt} + \epsilon^a_{jt})$. We then estimate λ

and ρ^a using GMM.[17] The moment conditions assume orthogonality between $(-\xi^a_{jt} + \epsilon^a_{jt})$ and the prices of coal, the lagged prices of coal, and the plant's current capacity. Denote the prices and their lagged values by Z^a_{jt}.

$$E[Z^a_{jt}(-\xi^a_{jt} + \epsilon^a_{jt})] = 0.$$

The input prices are valid instruments because they are not correlated with ξ^a_{jt} and ϵ^a_{jt}. First, ξ^a_{jt} contains the shock to FGD O&M cost which includes the cost of chemical feed, labor, waste disposal, and other costs. By construction, this shock is uncorrelated with the price of coal, lagged price of coal, and the current capacity. Second, since from equations (13) and (12), ω^a_{jt} does not affect the optimal (b_{jt}, s_{jt}) for a plant, ξ^a_{jt} also does not affect the price of coal for the plant. Lastly, ϵ^a_{jt} is assumed to be uncorrelated with the prices of coal and labor.

4.3 Estimation of the Total Variable Cost Function

To capture the potential consistency in the plant's generation efficiency, we assume that the current ω^y_{jt} affects ω^y_{jt+1}.

$$\omega^y_{jt+1} = g^y(\omega^y_{jt}) + \xi^y_{jt+1} = \rho^y_0 + \rho^y_1\omega^y_{jt} + \rho^y_2(\omega^y_{jt})^2 + \xi^y_{jt+1}, \tag{21}$$

where ξ^y_{jt+1} is the shock to the productivity. Because ω^y_{jt+1} measures the heterogeneity in generation efficiency due to either input quality or the overall operating efficiency of a plant, ξ^y_{jt+1} represents the shocks to the input quality or the overall efficiency. We assume that ξ^y_{jt+1} and ξ^y_{jt} are independent of each other because any persistency in the efficiency would be captured by the other terms already.

[17] Notice that the constants λ_0 and ρ^a_0 are not separately identified.

In this step, we use the total variable cost function to estimate the Cobb-Douglas coefficients in the electricity generation function, $\beta = (\beta_0, \beta_k, \beta_l, \beta_h)$, and the linear parameters in the generation productivity transition function, $\rho^y = (\rho_0^y, \rho_1^y, \rho_2^y)$. Rearranging the total cost functions for the two types of plants in equation (14) and (15) and taking logarithms, we get

$$\log \tilde{C}_{jt} = \log\left(\frac{w_{jt}^c(b_{jt}^*, s_{jt}^*; f_{jt}) + 2p_{jt}s_{jt}^*}{b_{jt}^*}\right) + \frac{1}{\beta_h}(\log y_{jt} - \beta_0 - \beta_k \log k_{jt} - \beta_l \log l_{jt} - \omega_{jt}^y) + \epsilon_{jt}^c, \tag{22}$$

where we use $\log h_{jt} = \frac{1}{\beta_h}(\log y_{jt} - \beta_0 - \beta_k \log k_{jt} - \beta_l \log l_{jt} - \omega_{jt}^y)$, and ϵ_{jt}^c is a plant-year specific cost shock. For both types of plants, the new cost term is $\tilde{C}_{jt} = w_{jt}^c n_{jt} + 2p_{jt}s_{jt}n_{jt}$, which is the sum of coal cost and permit cost if the plants uses permits to emit all the SO_2 generated.

Taking the logarithm of the lagged production function, we get

$$\omega_{jt-1}^y = \log y_{jt-1} - \beta_0 - \beta_k \log k_{jt-1} - \beta_l \log l_{jt-1} - \beta_h \log h_{jt-1}.$$

Plug this expression into the the transition function, $\omega_{jt}^y = g^y(\omega_{jt-1}^y) + \xi_{jt}^y$ and replace the ω_{jt}^y in equation (22) with the transition function. Equation (22) becomes

$$\begin{aligned}\log \tilde{C}_{jt} = \log\left(\frac{w_{jt}^c(b_{jt}^*, s_{jt}^*; f_{jt}) + 2p_{jt}s_{jt}^*}{b_{jt}^*}\right) + \frac{1}{\beta_h}(\log y_{jt} - \beta_0 - \beta_k \log k_{jt} - \beta_l \log l_{jt} \\ - g^y(\omega_{jt-1}^y)) - \frac{1}{\beta_h}\xi_{jt}^y + \epsilon_{jt}^c.\end{aligned} \tag{23}$$

We estimate the parameters in (23) using GMM, in which the moments conditions are based on the orthogonality between the composite error term and instrumental variables.

$$E\left[Z_{jt}^c\left(-\frac{1}{\beta_h}\xi_{jt}^y + \epsilon_{jt}^c\right)\right] = 0, \tag{24}$$

where Z^c_{jt} includes the logarithm of the current plant-year specific prices of coal and labor, their lagged values, and the lagged capacity and labor inputs.

We assume that the instruments are uncorrelated with ξ^y_{jt} and ϵ^c_{jt} for the following reasons. First, the FOCs for b_{jt} and s_{jt} imply that ω^y_{jt} does not affect (b_{jt}, s_{jt}), so ξ^y_{jt} does not affect the coal price or its lagged value. Second, ξ^y_{jt} represents the shock to heterogeneous input quality to the extent that such heterogeneity exists. However, we argue that heat and the average labor are quite homogeneous inputs across plants. Energy is measured in terms of thermal content, so there is also no omitted quality differential for heat input. The differences in average labor quality across plants are minimal due to the rigid production process. Thus, ξ^y_{jt} does not contain quality differentials of heat or labor across plants, which implies that this error is not correlated to the prices of coal and labor. Third, since labor and capital are assumed to be exogenous in the short run, ξ^y_{jt} does not affect the lagged labor and capacity inputs. Lastly, ϵ^c_{jt} is the idiosyncratic error in the total variable costs, which is uncorrelated with the input prices and lagged inputs.

5 Estimation Results

The estimation results of the coal price function are in Table 2. We compute the marginal prices of sulfur content and Btu content using equations (17) and (18). Table 3 shows the marginal prices. We find that the marginal price of b_{jt} is positive for all plants in all years, and the marginal price of sulfur content is negative for all plants in all years. The price per short ton of coal goes up by \$1.79 on average if b_{jt} increases by one million Btu per ton of coal. The standard error of the marginal price of b_{jt} is \$0.048. The coal price per short ton drops by \$1.89 if the sulfur content increases by 1%. The standard error of the marginal price of sulfur is \$0.11.

From Table 2, we find that, as s_{jt} and A_t increase, the marginal price of sulfur becomes less negative, implying that relaxing regulation leads to smaller impact of sulfur content on coal price. Coal price also decreases with the delivered quantity. The estimated coefficient for $log(n_{jt})$ of -0.061 implies that the average coal price elasticity with respect to quantity is -0.0015. Given this low elasticity, we assume that the quantity discounts do not affect the plants' choices of (b_{jt}, s_{jt}).

Table 2: Coal Price Function

Coefficient	Estimates
s_{jt}	0.935
	(0.824)
b_{jt}	2.200***
	(0.068)
s_{jt}^2	0.288***
	(0.037)
b_{jt}^2	-0.004***
	(0.000)
$s_{jt} * b_{jt}$	-0.181***
	(0.030)
$s_{jt} * allowances$	0.048***
	(0.007)
$log(n_{jt})$	-0.061***
	(0.022)

Parentheses contain estimated asymptotic standard errors.

The symbol *** indicates significance at the .01 level using a two-tailed t-test.

Table 3: Estimated Marginal Prices of Sulfur and Btu

	mean	std. err.	[95% Conf. Interval]	
$\frac{\partial w_{jt}^c}{\partial s_{jt}}$	-1.891	0.113	-2.112	-1.670
$\frac{\partial w_{jt}^c}{\partial b_{jt}}$	1.786	0.048	1.691	1.881

Table 4 shows the results from estimating the abatement cost function (20).[18] The estimate of λ is greater than one, which implies that the marginal abating cost of sulfur is increasing.[19] It also means that FGD equipment can reduce pollution control costs compared with using permits and emitting all the SO_2. This occurs because $\hat{\lambda} > 1$ implies that the marginal cost is lower than the permit prices up to a certain amount of sulfur. The estimates of ρ_1^a and ρ_2^a imply that $\frac{\partial \omega_{jt+1}^a}{\partial \omega_{jt}^a} = 0.91$, so there is significant persistency in a plant's abatement efficiency. Using these estimation results and the FOC for s_{jt}, we can compute the permit prices at the plant-year level.[20]

Table 4: Abatement Cost Function

	log(abatement cost)
λ	4.382
	(0.802)***
ρ_0^a	10.89
	(6.508)*
ρ_1^a	0.031
	(0.237)***
ρ_2^a	0.017
	(0.001)***
N	170

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 5 shows the estimation results of the total variable cost function and the productivity transition equation. Our derived total variable cost function includes the Cobb-

[18] We dropped four plants in the estimation due to obvious mistakes in the amount of coal, which leaves us a sample of 760 plant-year observations.

[19] This result is very consistent with that obtained by running through the Berkenpas, Rubin, and Zaremsky (2007) Integrated Emission Control Model (IECM). In the IECM model, we use 80 data points for different values of the Btu and sulfur content for coal, MW output levels, control levels, and regions of the US. The ranges of these values are representative of our data.

[20] See Appendix B for the results on the average permit prices by year. We also compare them with the permit prices from the permit auction data.

Douglas production function parameters. The first column shows the results of estimating the cost function. We label this column as Cost Minimization (CM), which means that firms minimize variable cost to produce exogenous amount of output. To compare our results with the production function estimation approaches which assume firms endogenously choose output, we present the results using the LP estimation method in the second column and the ACF estimation method in the third column.[21]

Using the cost function approach, the estimates of (β_k, β_h) in the CM column are positive and significant, with $\hat{\beta}_k = 0.185$, $\hat{\beta}_h = 0.905$. The impacts of labor on the total generation is -0.030 and insignificant. This is because the generation process of electricity in coal plants is only directly affected by capital and heat. Plants exhibit slight increasing returns in electricity generation on average, $\hat{\beta}_l + \hat{\beta}_k + \hat{\beta}_h = 1.06 > 1$. This is consistent with findings in the literature as summarized by Atkinson (2018). The estimate of ρ_1^y is 0.866 and significant, implying that lagged productivity significantly influences the current

[21]The LP approach assumes that capital is pre-determined and labor is variable in each period and uses a two-step estimation method. In the first step, the estimation equation is

$$y_{jt} = \beta_0 + \beta_k k_{jt} + \beta_l l_{jt} + \beta_h h_{jt} + g_t(k_{jt}, a_{jt}, h_{jt}) + \eta_{jt} = \beta_l l_{jt} + \phi_t(k_{jt}, a_{jt}, h_{jt}) + \eta_{jt},$$

where $g_t(k_{jt}, a_{jt}, h_{jt})$ is the inverted variable input function to get the unobserved ω_{jt}^y. We approximate the ϕ_t with the third-order polynomials and estimate the coefficients using GMM. The moment conditions are

$$E\left[\eta_{jt}(\theta_0) z_{1jt}\right] = 0,$$

where $z_{1jt} = (1, k_{jt}, a_{jt}, h_{jt})$ and higher order polynomials of them.

In the second step, the ω_{jt}^y is replaced by its transition function. The production function becomes

$$y_{jt} = \beta_0 + \beta_k k_{jt} + \beta_l l_{jt} + \beta_h h_{jt} + f(\omega_{jt-1}, x_{jt-1}) + \xi_{jt} + \eta_{jt},$$

in which the transition function is

$$f(\omega_{jt-1}, x_{jt-1}) = \rho_0 + \rho_1 \omega_{jt-1} + \rho_2 \omega_{jt-1}^2 + \rho_3 \omega_{jt-1}^3 + x'_{jt-1}\rho.$$

We estimate the parameters using GMM, and the moment conditions are

$$E\left[(\xi_{jt} + \eta_{jt})(\theta_0) z_{2jt}\right] = 0,$$

where $z_{2jt} = (1, k_{jt}, l_{jt-1}, a_{jt-1}, h_{jt-1})$ and higher order polynomials of them.

The ACF approach is similar to the LP approach but assumes that labor is also pre-determined. In the first step, the inverted intermediate input function also depends on labor. That is, productivity is replaced by $g_t(k_{jt}, l_{jt}, a_{jt}, h_{jt})$. The second step is similar to the LP method.

productivity. More productive plants are consistently more productive over time.

Table 5: Estimates of Parameters in the Production Function and the Transition Function

	(1) CM	(2) LP	(3) ACF
β_k	0.185	0.008	0.022
	(0.104)*	(0.021)	(0.020)
β_l	-0.030	-0.024	-0.038
	(0.029)	(0.018)	(0.012)***
β_h	0.905	1.057	1.052
	(0.097)***	(0.014)***	(0.012)***
ρ_1^y	0.866	0.846	0.954
	(0.016)***	(0.012)***	(0.032)***
N	760	760	760

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 5 shows that the LP and ACF approaches under-estimate β_k and over-estimates β_h. These biases are caused by that the production function estimation approaches assume endogenous outputs and estimate non-viable estimation equation when firms face exogenous outputs and choose inputs to minimize cost. As in equation (2), the heat input is a function of y_{jt}, so the inverted ω_{jt}^y is also a function of y_{jt}. Thus, in the first step estimation of LP and ACF, y_{jt} appears on both sides of the equation, which is a non-viable estimation equation.

The generation efficiency and the abatement efficiency are negatively correlated. The partial correlation is -0.23 after controlling for the logarithm of total generation. This implies that, to a moderate degree, more productive plants are less efficient in abating SO_2. We also find that the larger plants have lower generation efficiencies than smaller plants, with a partial correlation of -0.88. Both partial correlation coefficients are significant at less than the 0.01 level.

6 Plant Efficiencies and Unit Closures

From 2006-2017, in aggregate our sample plants closed 45% of their generating units. We estimate logit models to explain unit closing over this time period. Our first logit model is for all units (those with and without FGD scrubbers). Our second logit model is for only units with FGD devices. We assign 0 if the plant is closed and 1 if it is operating. For the first model, we employ as explanatory variables the percent of sulfur that is scrubbed in 2016, the year the plant was built, the centered fitted value of ω_{jt}^{y}, the centered fitted value of ω_{jt}, the weighted average price of coal in 2007, megawatt capacity of the plant, and the weighted average b_{jt} and s_{jt} in 2007. We chose the year 2007 since this gives plants enough lead time to find other generating sources if they close down coal-fired units in response to changes in w_{jt}, b_{jt}, and s_{jt}.

Results from the first model appear in column (1) of Table 6. They indicate that increasing the percent of SO_2 that is scrubbed by 1% point increases the probability of operating by about .32%, a one-year-newer plant is 1.05% more likely to operate, every 100 MW increase in capacity increases the probability of staying open by 4.74%, and reducing the average s_{jt} by 1% point increases the probability of survival by 8.68%. Since the fitted output productivity measure is not significant, we see no evidence that more productively efficient plants are more likely to survive. This may be due in part to court challenges and increasingly strict environmental standards that have shut down more efficient plants.[22]

[22] For example, two of the most efficient plants (highest ω_{jt}^{y} on average) are the Barry plant of Alabama Power Co. and Cholla of the Arizona Public Service Co. Three units of the Barry plant either closed or converted to natural gas units because of legal challenges by EPA. The Cholla plant closed several units because of the cost of installing scrubbers made it uneconomical to operate. http://www.pinnaclewest.com/newsroom/.

Table 6: Partial Effects on Probability of Plant Operating in 2018

VARIABLES	(1) ALL UNITS	(2) FGD UNITS
pct. scrubbed 2016	0.00317***	0.00826***
	(0.000910)	(0.00105)
year built	0.0105**	
	(0.00447)	
centered $\hat{\omega}^y$	-0.596	18.98***
	(0.379)	(2.187)
centered $\hat{\omega}^a$		0.339***
		(0.0308)
wtd. avg. w_{jt} 2007	-0.0321	
	(0.0675)	
MW Capacity	0.000474***	
	(0.000165)	
wtd. avg. b_{jt} 2007	-0.00990	-0.271***
	(0.0169)	(0.0239)
wtd. avg. s_{jt} 2007	-0.0868**	-1.029***
	(0.0432)	(0.115)
Observations	242	59

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

For the subsample of 59 units with FGD devices, we add the fitted value of ω^a_{jt} as an explanatory variable but drop year built, the weighted average w_{jt}, and MW capacity.[23] We find that increasing the percent of coal scrubbed by 1% point increases the probability of survival by .83%, increasing production efficiency by 1% point increases the probability of survival by about 19%, increasing abatement efficiency by 1% point increases the probability of operating by about .34%, a 1% increase in the weighted-average b_{jt} of coal reduces the probability of operating by .27%, while a 1% increase in the weighted-average s_{jt} of coal reduces the probability of operating by approximately 1.03%.[24] By contrast with the

[23]Due to our small sample size with the second model, to achieve convergence, we dropped variables with weaker priors for inclusion.e have provided stronger theoretical reasons to include the other variables.

[24]The magnitude of ω^a is much smaller because it has an exponential impact on the abatement cost. A one point increase in ω^a increases the abatement cost by 1.40 times.

results in column 1, the effect of $\hat{\omega}^y$ on the probability of operating is much greater for FGD units than for all units. The positive coefficient for ω^a is consistent with the goal of the Acid Rain Program, whose aim was to increase abatement efficiency through permit trading.[25] Increasing the percent of coal scrubbed as well as switching to lower b_{jt} and lower s_{jt} affect the probability of survival in the same direction but to a considerably larger extent than for all plants taken together.

7 Counterfactuals

7.1 Coal Cost Subsidy

In January 2018 Department of Energy Secretary Rick Perry proposed subsidizing coal plants by paying for 90 days' worth of coal on-site.[26] This payment would be equivalent to a subsidy equal to 25% of the cost of annual coal consumption. To analyze the impact of this subsidy on the coal plants' choices of coal, we use the structural model and the estimation results to compute the new cost minimizing Btu content and sulfur content for each plant in each year. Denote the subsidy by d, so $d = 25\%$. If plant j chooses (b_{jt}, s_{jt}) in year t, the subsidized coal price is

$$\tilde{w}^c_{jt}(b_{jt}, s_{jt}, f_{jt}) = (1-d)\left[\bar{w}^c_{jt}(b_{jt}, s_{jt}) + w_{jt}(f_{jt})\right],$$

[25]See the Environmental Protection Agency web site: https://www.epa.gov/airmarkets/acid-rain-program

[26]For more details see https://www.nytimes.com/2018/01/08/climate/trump-coal-nuclear.html. In a major blow to the Trump Administration's efforts to revive America's declining coal industry and specifically higher-sulfur coal, this proposal was rejected by the Federal Energy Regulatory Commission.

where $\tilde{w}^c$ represents the subsidized price. The cost minimization problem in year t becomes

$$\begin{cases} \min_{b_{jt},s_{jt}} \tilde{w}^c_{jt}(b_{jt}, s_{jt})n(b_{jt}; X_{jt}, \omega^y_{jt}) + C^s(s_{jt}, b_{jt}; X_{jt}, \omega^y_{jt}, \omega^a_{jt}), & \text{if FGD, or} \\ \min_{b_{jt},s_{jt}} \tilde{w}^c_{jt}(b_{jt}, s_{jt})n(b_{jt}; X_{jt}, \omega^y_{jt}) + C^s(s_{jt}, b_{jt}; X_{jt}, \omega^y_{jt}), & \text{if non-FGD.} \end{cases}$$

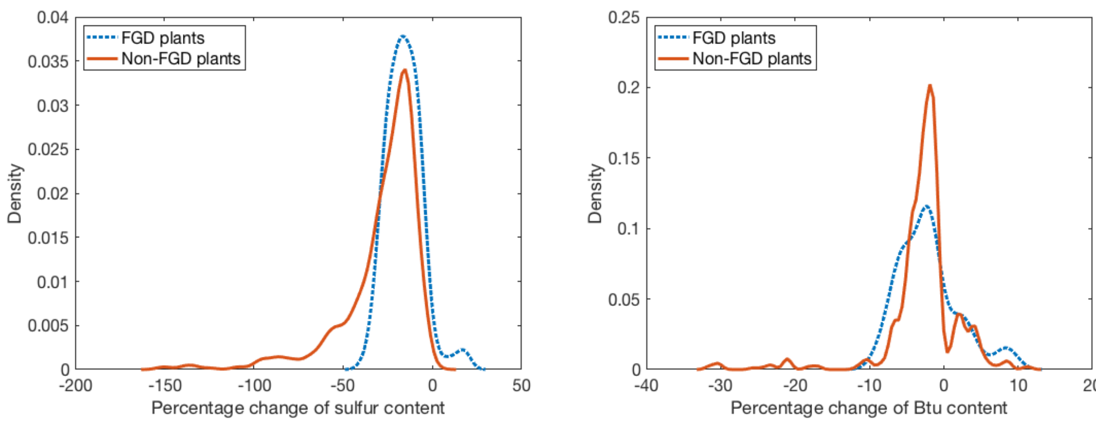
To solve for the new cost minimizing (b_{jt}, s_{jt}), we assume that a plant can only adjust (b_{jt}, s_{jt}) up to two standard deviations of the plant's Btu content and sulfur content in the data. For plant j in year t, the range for the new Btu content is $[b_{jt} - 2\sigma^b_j, b_{jt} + 2\sigma^b_j]$, where b_{jt} is the plant's Btu content in the data, and σ^b_j is the standard deviation of plant j's Btu content from 1995-2005. The average value of σ^b_j over all plants is 0.43 million Btu per ton, which is about 1.93% of the average Btu content. Similarly, the range for the new sulfur content is $[s_{jt} - 2\sigma^s_j, s_{jt} + 2\sigma^s_j]$, where s_{jt} is the plant's Btu content in the data, and σ^s_j is the standard deviation of plant j's sulfur content from 1995-2005. The average value of σ^s_{jt} over all plants is 1.4%, which is 13.13% of the average sulfur content. These ranges cover 95% of the observed ranges in the data.

Since we do not model the demand and the supply in the permit trading market, we assume that the plant-year specific permit prices are the same as in the data. We also assume that the shipping charge stays the same when a plant chooses a different pair of (b_{jt}, s_{jt}) because we do not have a one-to-one mapping from (b_{jt}, s_{jt}) to a mine and the distance between a mine and a plant. However, this assumption is not very restrictive because a plant can buy coal with very different Btu content and sulfur content from adjacent mines.[27]

The plants face trade-offs when adjusting b_{jt} and s_{jt}. Decreasing b_{jt} has two opposite impacts on the total variable cost. On one hand, the coal price is lower when b_{jt} decreases, so lowering b_{jt} can reduce total expenditures on a given amount of coal. On the other

[27]See footnote 11 in Section 14 for more details.

hand, the plants need to buy more coal to maintain the same level of total heat input. This increases the coal cost and the pollution control cost for a given s_{jt}. Decreasing s_{jt} also has two opposite impacts on the total variable cost. First, the coal price increases when s_{jt} decreases, so the total cost of coal increases for a given b_{jt}. Second, the pollution control cost decreases due to lower s_{jt}. Therefore, how each plant would adjust the b_{jt} and s_{jt} depends on the relative strength of the impacts on the coal cost and the pollution control cost.



(a) (b)

Figure 4: Density of percentage changes in s_{jt} and b_{jt} with the coal subsidy

Figure 4a shows the densities of the percentage change in sulfur content at the plant-year level. The dotted curve is for FGD plants and solid curve is for non-FGD plants. Most plants would lower the sulfur content in most years, but in a few years, non-FGD plants reduce sulfur content more than FGD plants. On average, the FGD plants and non-FGD plants would decrease the sulfur content by 15.97% and 28.22%, respectively. Figure 4b shows the densities of the percentage change in Btu content. The dotted curve is for FGD plants and the solid curve is for non-FGD plants. We find that majority of the plants would lower the Btu content in most years. On average, the FGD plants and non-FGD plants would reduce Btu content by 2.08% and 2.72%, respectively.

Table 7 summarizes the impacts of the subsidy on the plants' abatement level, emission, and costs. The numbers are the averages over the plant-year observations. We find that for FGD plants, the average coal price would decrease by 27.85%. Due to lower b_{jt}, the amount of coal would increase to maintain the same level of total heat input. On average, the FGD plants would buy 2.28% more coal. The amount of coal abated would increase by 14.29% on average. The percentage increase of the abatement level is greater than that of the coal quantity because of the lower sulfur content. The total amount of sulfur would decrease by 14.18% due to the significantly lower sulfur content. The FGD plants would save 26.41% on the total cost of coal including the 25% coal-cost subsidy, the 26.45% reduction in the total cost of permits, and the 16.85% reduction in abatement costs. Their total variable costs would decrease by 26.94%. The average subsidy for an FGD plant per year would be \$18.17 million, which leads to a total subsidy of \$3.60 billion for the 18 FGD plants during the 11 sample years.

Table 7: Impacts of a 25% Coal Cost Subsidy

	FGD			Non FGD		
	data	subsidy	change(%)	data	subsidy	change(%)
Sulfur content (%)	1.90	1.62	-15.97%	0.92	0.68	-28.22%
Btu content (10^6/ton)	21.61	21.18	-2.08%	22.98	22.39	-2.72%
Coal price (\$)	24.34	17.81	-27.85%	33.04	24.30	-27.25%
Coal quantity (10^6 tons)	4.06	4.15	+2.28%	2.40	2.49	+3.13%
Coal abatement (10^6 tons)	3.21	3.55	+14.29%			
Sulfur (10^4 tons)	7.64	6.69	-14.18%	1.95	1.48	-26.42%
SO_2 emission (10^4 tons)	3.31	2.26	-26.45%	3.90	2.96	-26.42%
Coal costs (\$$10^6$)	98.07	72.68	-26.41%	76.08	56.85	-25.35%
Permit costs (\$$10^6$)	1.79	1.28	-26.45%	3.77	2.87	-26.42%
Abatement costs (\$$10^6$)	7.96	5.42	-16.85%			
Total costs (\$$10^6$)	107.82	79.38	-26.94%	79.85	59.72	-25.36%
Subsidy (\$$10^6$)		18.17			14.21	

The right panel of Table 7 shows the impact of the 25% coal subsidy on the non-FGD plants. Their average coal price would decrease by 27.25%. The quantity of coal would increase by 3.13% due to the lower Btu content. The total amount of sulfur in the purchased coal would decrease by 26.42%. The non-FGD plants would save 25.35% on the total coal costs and 26.42% on the permit costs. Their total variable costs would decrease by 25.36%. The average subsidy for a non-FGD plant per year would be $14.21 million, which implies a total subsidy of $9.38 billion for the 58 non-FGD plants during the 11 years.

7.2 An Increase in the Emission Allowances

The Trump Administration has consistently opposed strengthening air quality regulations promulgated under the Clean Air Act and has even rolled back many regulations. This Administration has proposed or implemented less stringent standards on power plants.[28] Chief among them was the repeal of stricter limits on SO_2 emissions from existing coal-fired power plants, thereby potentially stimulating the demand for coal by power plants. Under the Obama Adiministration's Clean Power Plan, EPA projects a reduction in SO_2 of 24% by 2030 compared with 2005 levels.[29] However, under the Trump Administration's proposed plan, the EPA projects that SO_2 would be cut by only 1-2%. Based on these recent rollbacks of regulatory standards, we are interested in the impacts of increased SO_2 emissions on the Btu and sulfur choices of coal-fired power plants.

We simulate this type of action by modeling an increase in free allowances for all plants. In this counterfactual case, we consider the scenario in which the permit trading system is non-functional and all plants receive free allowances that are 120% of their current SO_2 emission level. We assume non-functionality of the current permit trading market

[28] See https://www.vox.com/energy-and-environment/2018/1/26/16936104/epa-trump-toxic-air-pollution and "76 Environmental Rules on Way Out under Trump": https://www.nytimes.com/interactive/2017/10/05/climate/trump-environment-rules-reversed.html.

[29] See https://www.washingtonpost.com/national/health-science/new-trump-power-plant-plan-would-release-hundreds-of-millions-of-tons-of-co2-into-the-air/2018/08/18/be823078-a28e-11e8-83d2-70203b8d7b44_story.html?noredirect=on&utm_term=.816359192c84f.

since changes in the regulations on SO_2 permit trading made the market inactive, with effectively zero trades and zero permit prices since 2010. This was due primarily to a D.C. Appeals Court ruling that year which stuck down the Clean Air Interstate Rule (CAIR), the increased use of FGD devices, and the reduced use of coal for generation.[30] Thus, for this counterfactual, we model the increase in SO_2 permits in a non-functional permit market during our sample period.

With the free allowances, the FGD plants would only abate the SO_2 that exceeds the free allowances. The new cost minimization problem for an FGD plant j in year t is to choose (b_{jt}, s_{jt}) to minimize the cost of coal generation and SO_2 abatement.

$$\min_{b_{jt},s_{jt}} w^c_{jt}(b_{jt}, s_{jt})n(b_{jt}; X_{jt}, \omega^y_{jt}) + e^{\lambda_0 - \omega^a_{jt}}(n^a_{jt}s_{jt}r_{jt})^\lambda, \tag{25}$$

subject to the emission constraint that

$$2s_{jt} * (n(b_{jt}; X_{jt}, \omega^y_{jt}) - n^a(b_{jt}, s_{jt}; X_{jt}, \omega^y_{jt}, \omega^a_{jt})r_{jt}) \leq 1.2 * S^{e,data}_{jt}. \tag{26}$$

The non-FGD plants would have to cap the total sulfur input and SO_2 output to comply with the new emission regulation because they do not have the technology to abate any excessive SO_2. For a plant without FGD, the cost minimization problem is

$$\min_{b_{jt},s_{jt}} w^c_{jt}(b_{jt}, s_{jt})n(b_{jt}; X_{jt}, \omega^y_{jt}), \tag{27}$$

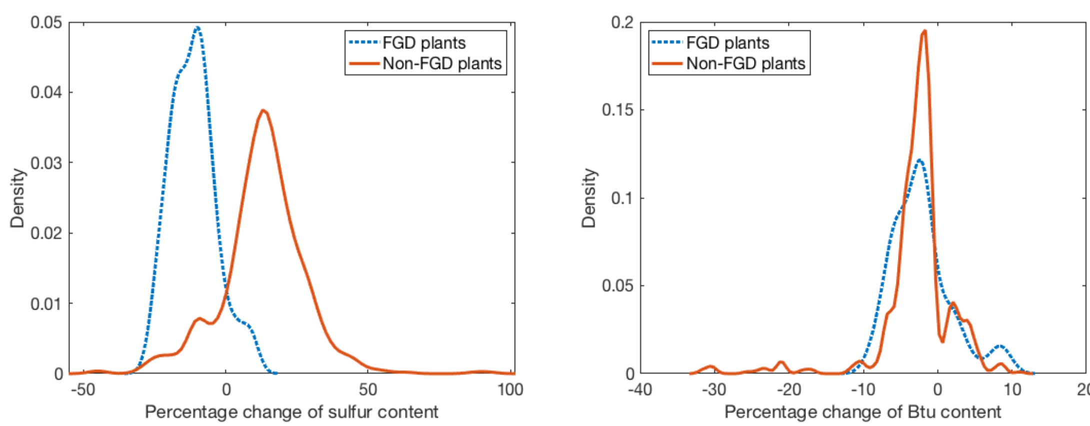
subject to the emission constraint that

$$2s_{jt} * n(b_{jt}; X_{jt}, \omega^y_{jt}) \leq 1.2 * S^{e,data}_{jt}. \tag{28}$$

[30] In striking down CAIR, the court ruled that banked allowances as well as allowances awarded to states under CAIR were disallowed. Interstate trades were also disallowed. See https://www.eia.gov/todayinenergy/detail.php?id=4830

The emission constraints for both types of plants would be binding when plants minimize costs because the permit system is non-functional. This implies that plants cannot sell unused permits.

As in the previous counterfactual analysis, we impose constraints on the (b_{jt}, s_{jt}) such that the plants can only adjust them up to two standard deviations in each direction. Figure 5a shows the density of the percentage change in sulfur content at the plant-year level. The blue curve is for FGD plants and red curve is for non-FGD plants. We find that the FGD plants would lower the sulfur content by 11.43% on average, while the non-FGD plants would increase it by 13.01%. The FGD plants can reduce abatement costs by lowering s_{jt}, but non-FGD plants do not have abatement costs to displace. Figure 5b shows the density of the percentage change in Btu content. The blue curve is for FGD plants and the red curve is for non-FGD plants. We find that most of the plants would lower their Btu content. On average, the FGD plants and non-FGD plants would reduce Btu content by 2.23% and 2.72%, respectively.



(a) (b)

Figure 5: Density of percentage changes in s_{jt} and b_{jt} with a 20% allowance increase

The increase in non-tradable allowances has two direct and one indirect impacts on s_{jt}.

The first direct impact is that the plants have an incentive to choose a higher s_{jt} for a fixed b_{jt} because this would lower the coal price and the plants can cover the additional SO_2 with additional allowances. The second direct effect is that FGD plants could choose a lower s_{jt} so that they can significantly reduce abatement costs. This effect does not apply to non-FGD plant as they rely on permits and allowances exclusively to comply with environmental regulations. The indirect impact arises through b_{jt}. If a plant lowers b_{jt}, then the total quantity of coal increases, so the plant would choose a lower s_{jt} to comply with the emission regulation. The combination of these effects determines the new s_{jt}. We find the FGD plants would reduce s_{jt} due to the allowance increase, while the non-FGD plants would increase s_{jt}.

For both types of plants, the allowance increase and higher s_{jt} can raise or lower b_{jt}. If the allowance increase also raises b_{jt}, then they pay higher prices for coal and buy less coal to reach the same amount of heat input. In this case, the total sulfur may increase or decrease, and the plants may waste the free allowances if the total sulfur does not increase by 20%. If they lower b_{jt}, then they pay lower prices for coal and buy more coal. In this case, the total amount of sulfur will increase due to higher s_{jt} and more coal purchases, and they can use the additional free allowances for the increase in sulfur. The results show that both types of plants will lower b_{jt} to minimize the cost of coal.

Table 8 shows the impact of the 20% emission allowance increase on coal price, abatement level, and costs for FGD and non-FGD plants. The results are averages over all years. For FGD plants, the price of coal would decrease by 4.29% and the quantity of coal would increase by 2.44%. The total cost of coal would decrease by 2.22%. The amount of coal abated by FGD plants would decrease by 19.02%. The total sulfur in coal would decrease by 9.29%. Due to lower total sulfur and more allowances, the abatement cost would decrease dramatically, by 67.38%. Lastly, the total costs of coal and abatement would decrease by 7.48%.

For the non-FGD plants, the coal price would decrease by 5.06% and the quantity of coal would increase by 3.12%. The total cost of coal would decrease by 2.59%, while the total sulfur in the coal would increase by 16.22%. This change is less than 20% due to the constraints on s_{jt}. The non-FGD plants could reduce their coal and abatement costs by 2.69% on average, considerably less than that for the FGD plants.

Table 8: Impacts of a 20% SO_2 Allowance Increase

	FGD			Non FGD		
	data	allowance	change	data	allowance	change
Sulfur content (%)	1.90	1.66	-11.43%	0.92	1.05	+13.01%
Btu content (10^6/ton)	21.61	21.15	-2.23%	22.98	22.39	-2.72%
Coal price ($)	24.34	23.63	-4.29%	33.04	31.70	-5.06%
Coal quantity (10^6 tons)	4.06	4.15	+2.44%	2.40	2.49	+3.12%
Coal abatement (10^6 tons)	3.21	2.79	-19.02%			
Sulfur (10^4 tons)	7.64	6.87	-9.29%	1.95	2.25	+16.22%
Coal costs ($$10^6$)	98.07	96.64	-2.22%	76.08	74.38	-2.59%
Abatement costs ($$10^6$)	7.96	2.87	-67.38%			
Coal & Abatement Costs ($$10^6$)	106.03	99.52	-7.48%	76.08	74.38	-2.69%

8 Conclusions

The traditional approach of OP/LP/ACF to production function estimation requires the assumption of unconstrained profit maximization which may be inappropriate for many firms and plants who face exogenous output targets. This approach would lead to a non-viable estimation equation in which output is a function of itself. To solve this problem, we model the cost minimization problem of firms and derive the total variable cost function which depends on the production function parameters. This cost function and a transition function of the unobserved productivity together provide identification for the parameters. Our cost function approach contributes to the production function estimation literature by relaxing the assumption of unconstrained profit maximization with endogenous outputs.

We apply our approach to the US coal-fired power plants which minimize variable costs to generate electricity. Due to rate-of-return determined prices for regulated utilities and the least-cost dispatch methods used by power pools to which most deregulated utilities belong, utility plants face exogenous output targets. In fact, most of the empirical studies of the electric utility sector assume output-constrained cost minimization by plants. In our model, a plant's total variable cost includes both the cost of coal and the pollution control costs. For FGD plants, the pollution control costs include the sulfur dioxide abatement cost and the permit cost for emission. For non-FGD plants, the pollution control costs are simply the permit costs. Plants choose the sulfur content and Btu content of coal to minimize the total variable cost. We consider plant heterogeneity in both the unobserved generation efficiency and unobserved abatement efficiency.

Using data on the 80 largest plants from 1995-2005, we estimate the coal price as a function of its characteristics, the abatement cost function of SO_2, and the production function parameters via our cost function approach. In the coal price estimation, we control for confidential shipping charges using plant-mine fixed effects and find negative implicit prices of sulfur and positive implicit prices of Btu. The abatement cost function estimation results show that the plants have increasing marginal abatement cost, which is consistent with engineering estimates. Estimated production function parameters indicate moderate increasing returns to scale. We compare our results with those using the LP/ACF approach and find substantial differences.

To our knowledge, previous studies have not estimated implicit plant-year specific prices of sulfur and Btu that are free of confidential transportation charges. We find that on average, the coal price decreases by \$1.89 per ton if the sulfur content increases by 1%, and the coal price increases by \$1.79 per ton if the Btu content increases by a million Btu per ton of coal. We also provide the first specification and estimation of separate productivities for generation and abatement.

Using our estimates of generation and abatement productivities, along with other variables, we predict plant closings in 2017. For all plants, a lower sulfur content of fuel significantly increases the probability of remaining open. For FGD plants a lower Btu content and lower sulfur content increase the probability of plant survival. The most important of these variables is clearly the sulfur content.

Counterfactual analyses consider the implications of subsidies for coal consumption and an increase in permit allowances. Both of these measures mimic proposed and implemented policies of the Trump Administration which are designed to revitalize the coal industry by stemming reductions in the degradation of air quality. Ironically, the first measure substantially reduces sulfur content and SO_2 emissions from both FGD and non-FGD units. The second measure increases the sulfur content for non-FGD units and reduces it for FGD units. The latter result is counterintuitive but consistent with our model. Both measures moderately reduce the Btu content.

References

Ackerberg, D. A., Caves, K., Frazer, G., 2015. Identification properties of recent production function estimators Econometrica 83 (6), 2411-2451.

Atkinson, S. E., 2018. Panel Data in Energy Economics. Forthcoming in: Panel Data Econometrics Volume 2: Empirical Applications, ed. Mike Tsionas, Elsevier Publishing Co.

Atkinson, S.E., Dorfman, J.H., 2005. Bayesian measurement of productivity and efficiency in the presence of undesirable outputs: crediting electric utilities for reducing Air pollution, Journal of Econometrics 126, 445-68.

Atkinson, S. E., Tsionas, M. G., 2016. Directional distance functions: optimal endogenous directions, Journal of Econometrics 190 (2), 301-314.

Atkinson, S. E., Primont, D., Tsionas, M.G., 2018. Statistical inference in efficient production with bad inputs and outputs using latent prices and optimal directions. Journal of Econometrics 204 (2), 131-146.

Baltagi, B. H., Griffin, J.M., 1988. A general index of technical change, Journal of Political Economy 96 (1), 20-41.

Berkenpas, M. B., Rubin, E.S., and Zaremsky, C. J., 2007. User Manual: Integrated Environmental Control Model, Carnegie Melon University.

Bushnell, J. B., Mansur, E. T., Saravia, C., 2008. Vertical arrangements, market structure, and competition: An analysis of restructured US electricity markets. American Economic Review 98 (1), 237-266.

Christensen, L. R., Greene, W. H., 1976. Economies of scale in U.S. electric power of Political Economy, 84 (4), 655-76.

Carlson, C. Burtraw, D., Cropper, M., Palmer, K.L., 2000. Sulfur Dioxide Control by Electric Utilities: What Are the Gains from Trade? The Journal of Political Economy 108 (6), 1292-1326.

Chan, H.R., Fell, H., Lange, I., Li. S., 2017. Efficiency and environmental impacts of electricity restructuring on coal-fired power plants. Journal of Environmental Economics and Management 81, 1-18.

Department of Energy, Energy Information Administration (EIA) Form 423, 1995-2005, Washington, D.C.

Department of Energy. Energy Information Administration (EIA), 1999. U.S. Coal Reserves: 1997 Update, DOE/EIA-0529(97) Distribution Category UC-950.

Doraszelski, U., Jaumandreu, J., 2013. R&D and Productivity: Estimating Endogenous Productivity. Review of Economic Studies 80, 1338-1383.

Färe, R., Grosskopf, S., Noh, D., Weber, W., 2005. Characteristics of a polluting technology: theory and practice. Journal of Econometrics 126, 469-492.

Gandhi, A., Navarro, S., Rivers, D., 2017. On the Identification of Gross Output Production Functions. Working paper, Department of Economics, Madison, WI.

Gollop, F. M., Roberts, M. J., 1981. The sources of economic growth in the US electric power industry. In: Cowing, T. G. and Stevenson, R.E. (Eds.), Productivity Measurement in Regulated Industries. Academic Press, New York, 1981.

Gollop, F. M., Roberts, M. J., 1983. Environmental regulations and productivity growth: the case of fossil-fueled electric power generation. Journal of Political Economy, 91 (4), 654-74

Gollop, F.M., Roberts, M.J., 1985. Cost-minimizing regulation of sulfur emissions: regional gains in electric power. Review of Economics and Statistics 67 (1), 81–90.

Griliches, Z., Mairesse, J., 1998. Production functions: the search for identification. In: Strøm, S., (Ed.) Econometrics and Economic Theory in the 20th century. Cambridge University Press, pp. 169-203.

Kleit, A.N., Terrell, D., 2001. Measuring Potential Efficiency Gains from Deregulation of Electricity Generation: A Bayesian Approach. Review of Economics and Statistics 83 (3), 523-530.

Knittel, C.R., 2002. Alternative regulatory methods and firm efficiency: stochastic frontier evidence from the U.S. electricity industry. Review of Economics and Statistics 84 (3), 530-540.

Levinsohn, J., Petrin, A., 2003. Estimating production functions using inputs to control for unobservables, Review of Economic Studies 70 (2), 317-342.

Moody's Public Utility Manual, Moody's Investors Service, 1988-2005, New York.

Olley, G. S., Pakes, A., 1996. The dynamics of productivity in the telecommunications equipment industry. Econometrica 64 (6), 1263-1297.

Wooldridge, J. M., 2009. On estimating firm-level production functions using proxy variables to control for unobservables. Economics Letters 104, 112-114.

Appendix A

Table A.1. Plants and Firms

PLANT	FIRM	PLANT	FIRM
Barry	Alabama Power Co	Riverbend	Duke Energy Corp
Gorgas	Alabama Power Co	Muskingum River	Ohio Power Co
Colbert	Tennessee Valley Authority	W S Lee	Duke Energy Corp
Widows Creek	Tennessee Valley Authority	McMeekin	South Carolina Electric&Gas Co
Cholla	Arizona Public Service Co	Wateree	South Carolina Electric&Gas Co
Cherokee	Public Service Co of Colorado	Williams	South Carolina Electric and Gas
Comanche	Public Service Co of Colorado	Bull Run	Tennessee Valley Authority
Valmont	Public Service Co of Colorado	Cumberland	Tennessee Valley Authority
Lansing Smith	Gulf Power Co	Gallatin	Tennessee Valley Authority
Bowen	Georgia Power Co	John Sevier	Tennessee Valley Authority
Hammond	Georgia Power Co	Johnsonville	Tennessee Valley Authority
Mitchell	Georgia Power Co	Kingston	Tennessee Valley Authority
Joppa Steam	Electric Energy Inc	Carbon	PacifiCorp
Tanners Creek	Indiana Michigan Power Co	Clinch River	Appalachian Power Co
Bailly	Northern Indiana Pub Serv Co	Glen Lyn	Appalachian Power Co
Cayuga	PSI Energy Inc	Bremo Bluff	Virginia Electric & Power Co
R Gallagher	PSI Energy Inc	Chesterfield	Virginia Electric & Power Co
F B Culley	Southern Indiana Gas & Elec Co	Chesapeake	Virginia Electric & Power Co
Kapp	Interstate Power	John E Amos	Appalachian Power Co
Riverside	MidAmerican Energy Co	Kanawha River	Appalachian Power Co
LaCygne	Kansas City Power & Light Co	Philip Sporn	Central Operating Co
Big Sandy	Kentucky Power Co	Rivesville	Monongahela Power Co
E W Brown	Kentucky Utilities Co	Mt Storm	Virginia Electric & Power Co
Ghent	Kentucky Utilities Co	Pulliam	Wisconsin Public Service Corp
Green River	Kentucky Utilities Co	Weston	Wisconsin Public Service Corp
Cane Run	Louisville Gas & Electric Co	Dave Johnston	PacifiCorp
Mill Creek	Louisville Gas & Electric Co	Naughton	PacifiCorp
Paradise	Tennessee Valley Authority	James H Miller Jr	Alabama Power Co
Shawnee	Tennessee Valley Authority	R M Schahfer	Northern Indiana Pub Serv Co
Monroe	Detroit Edison Co	A B Brown	Southern Indiana Gas & Elec Co
St Clair	Detroit Edison Co	Welsh	Southwestern Electric Power Co
High Bridge	Northern States Power Co	Harrington	Southwestern Public Service Co
Asheville	Carolina Power & Light Co	Tolk	Southwestern Public Service Co
Lee	Carolina Power & Light Co	Pawnee	Public Service Co of Colorado
L V Sutton	Carolina Power & Light Co	Mountaineer	Appalachian Power Co
G G Allen	Duke Energy Corp	Belews Creek	Duke Energy Corp
Buck	Duke Energy Corp	Jim Bridger	PacifiCorp
Cliffside	Duke Energy Corp	Huntington	PacifiCorp
Dan River	Duke Energy Corp	Gen J M Gavin	Ohio Power Co
Marshall	Duke Energy Corp	North Valmy	Sierra Pacific Power Co

Appendix B

Using the estimates of the marginal prices of sulfur content and the FOCs with respect to the sulfur content, we compute the plant-year specific permit prices. Table B.1 shows the average estimated permit price and the EPA auction price of permits by year.

B.1. Plant-Year SO_2 Permit Prices ($/ton)

year	Estimates	EPA Auction Price
1995	105.55	130
1996	97.30	66
1997	93.11	107
1998	90.37	108
1999	86.60	201
2000	76.49	126
2001	80.31	174
2002	83.31	161
2003	83.43	172
2004	84.90	260
2005	86.71	690

Our estimates are close to the auction prices until 2001, when we begin to substantially underestimate actual prices. This occurs because from 2001 to 2005 a number of unforseen supply disruptions and speculation about the permanency of Executive Orders not captured in our model pushed permit price to very high levels. In 2001 the Clear Skies initiative was first announced. This was designed to cut SO_2 emissions by 73 percent from 2002 emissions of 11 million tons to a cap of 4.5 million tons in 2010, and 3 million tons in 2018. This rule never became law. In 2004 the Clean Air Interstate Rule (CAIR) was finalized. This rule in part required 28 Eastern states to make reductions in SO_2 of over 70% from 2003 levels. In 2005 the unforeseen events of train derailments of Powder River Basin coal shipments and disruptions of natural gas transmission caused by hurricanes Katrina and Rita pushed the average permit price to about $690. The CAIR initiative was vacated in

2008 after permit prices began to plunge from their peak in 2005 to current prices of less than one dollar.