Appendix For Online Publication

Dynamic Incentives and Permit Market Equilibrium in Cap-and-Trade Regulation

Yuta Toyama

A Online Appendix: Supplemental Analysis on Electricity Production Decision

In this Appendix, I provide a descriptive and simulation analysis on electricity production decisions. The purpose of these analyses is to examine the validity of exogenous production assumption in the structural model. Section 6.1 in the main body of the paper presents a summary of these analyses. Below, however, I provide their details. First, I conduct the difference-in-differences analysis to determine how the introduction of the Acid Rain Program affects the electricity generation in . Then, I run the regression analysis to estimate how the output prices respond to the regulation in A.2. Lastly, I examine the effect of the Acid Rain Program on the reallocation of production across generating units by constructing a merit-order curve in A.3.

A.1 Difference-in-differences Analysis on Electricity Generation

This section conducts a difference-in-difference (hereafter DID) regression to examine how the introduction of the Acid Rain Program affects the electricity production. To do this, I exploit the variation in the timing of the regulation across generating units. As was mentioned in Section 2.1, there are two groups of generating units: those regulated since 1995 (Group I units) and those regulated since 2000 (Group II units). Figure A1 shows the average capacity factor for each group each year.

While the data cover the period until 2003, all the units are treated (i.e., regulated) from 2000, implying that there are no control units after 2000. Following the practice in the recent DID literature (i.e., Goodman-Bacon 2021; Callaway and SantAnna 2021; Sun and Abraham 2021), I restrict the sample to the period between 1990 and 1999.[91]

I specify the regression equation as follows:

$$Y_{jm} = \alpha \text{GroupI}_j \cdot 1\{\text{after1995}\}_m + \beta' X_{jm} + u_j + u_m + u_{jm},$$

where Y_{jm} is the outcome variable of unit j in year-month m. I use as the main outcome variable capacity factor defined by $cf_{jm} = q_{jm}/k_j$, where q_{jm} is the net generation and k_j is

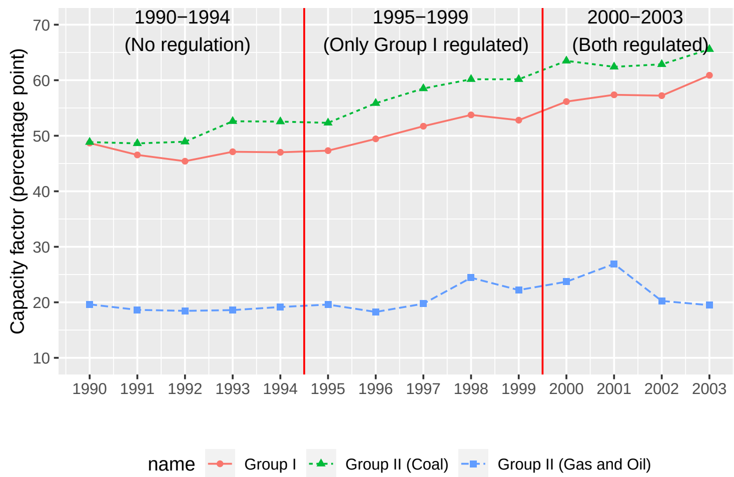
[91] Using the terminology of the recent DID literature, there are no "never-treated" units in the sample. The difference-in-difference regression using the sample after 2000 compares the Group II units with the Group I units, which are "already treated" units and thus are not adequate control units for the Group II units. See Goodman-Bacon 2021; Callaway and SantAnna 2021; Sun and Abraham 2021 for the details.

the nameplate capacity. GroupI_j is the dummy variable for the Group I units. $1\{\text{after1995}\}_m$ is the dummy variable that indicates the periods after 1995 (the beginning of Phase I). X_{jm} includes control variables such as the state-level electricity demand. Unit and time fixed effects are captured by u_j and u_m, respectively.

It is worth explaining the interpretation of the DID estimate of α. To estimate it, I compare the change in the utilization rate of the Group I units before and after 1995 and the change of the Group II units. Once the regulation is introduced in 1995, the Group I units have a higher marginal cost of generation due to the opportunity cost of SO_2 emissions under cap-and-trade. Such a cost increase might induce lower utilization of the Group I units. However, the Group II units could also change their electricity production via production reallocation. Specifically, the firm may reallocate the electricity production from Group I to Group II. Therefore, the DID estimate of α should be interpreted as an upper bound of the regulation effect on the Group I units.

The regression results are shown in Table A1. The estimates suggest that introducing the ARP decreases the capacity factor of group I units by 0.6–3.8 percentage points, depending on the choice of units in the control group. Although the effects are statistically significant, as shown in column (1), the economic significance is limited. Because the mean of the capacity factor falls within the range of 50–60 percentage points in the sample, electricity generation fell by at most 7.6% after the introduction of the cap-and-trade program. As shown in Section 2.3.1, this magnitude cannot account for the significant decrease in SO_2 emissions in the sample period. Instead, the adjustment of emissions rate of fuel, as discussed in Section 2.3.2, is the key channel of SO_2 abatement. The difference-in-differences analysis confirms that the abatement of SO_2 emissions was achieved primarily through adjusting emissions rates of fuel. This finding supports my modeling assumption of exogenous electricity production.

Figure A1: Trend of Capacity Factor of Group I and Group II units



Notes: The figure shows the trend of the capacity factor, defined by the ratio of net generation (output) to generation capacity, over time. I calculate the mean of the monthly-level capacity factor in each year for three groups: units regulated since 1995 (denoted as Group I), coal units regulated since 2000 (denoted as Group II (Coal)), and gas and oil units regulated since 2000 (denoted as Group II (Gas and Oil)).

Table A1: Difference-in-differences Regression of Capacity Factor

	Dependent variable:	
	Capacity factor	
	(1)	(2)
Treatment	−3.823	−0.643
	(0.633)	(0.568)
log(Electricity Demand))	38.481	42.705
	(1.698)	(1.182)
Control group	Coal only	Coal, Gas, and Oil
Observations	127,200	201,960
Adjusted R^2	0.527	0.656

Notes: Unit-level dummies and year-and-month dummies are included. Standard errors are clustered at the unit level.

A.2 Effects of Acid Rain Program on Output Prices

In this subsection, I examine how output prices change due to the introduction of the Acid Rain Program. Specifically, I run the following panel regression:

$$\log P_{st} = \theta Z_{st} + u_s + u_t + u_{st},$$

where P_{st} is the output price in state s in month-year t and Z_{st} is a measure of exposure to the Acid Rain Program. The measure of output price is defined as the average price for industrial/residential retail customers obtained from the form Form EIA-826 "Monthly Electric Utility Sales and Revenue Report with State Distributions" Energy Information Administration (1990–2003). For Z_{st}, I use (1) the number of generating units affected by the Acid Rain Program in state s and period t, (2) the total generation capacity of units that are regulated under the Acid Rain Program, and (3) the dummy variable indicating whether there exists at least one regulated unit in state s and period t. I use the data from 1990 to 2003. Note that there is both cross-sectional and temporal variation in the exposure to the Acid Rain Program.

Table A2 presents the estimation results. All the columns show that the impact on output price is both statistically and economically insignificant, which suggests that the introduction of the Acid Rain Program seems to have a limited impact on output prices.

A.3 Reallocation of Production across Generating Units based on a Merit-order Curve

I examine how the electricity generation would change if I consider the increase in the marginal cost of generation due to cap-and-trade. While the model assumes the exogenous electricity production, the cap-and-trade regulation might change the marginal costs of generation by incorporating the shadow cost of SO_2 emissions and thus the allocation of production across generating units. To investigate this point, I construct a merit-order curve based on the marginal cost and the capacity of generating units as well as examining how the shadow cost of SO_2 emissions changes the production pattern (e.g., Borenstein, Bushnell, and Wolak (2002), Bushnell, Mansur, and Saravia (2008), and Asker, Collard-Wexler, and De Loecker (2019)). I first introduce a model of electricity generation based on a merit-order curve in A.3.1 and show the summary statistics of fuel and permit costs in A.3.2. Using the model, I analyze the reallocation of electricity generation due to the introduction of the cap-and-trade in A.3.3. Lastly, I examine the extent to which coal units are likely to be a marginal unit in A.3.4.

Table A2: Effects of Exposure to the Acid Rain Program on State-level Electricity Price

	Dependent variable:		
	log (output price)		
	(1)	(2)	(3)
log(number of regulated units + 1)	−0.005		
	(0.008)		
log(capacity of regulated units + 1)		0.0001	
		(0.002)	
1at least one regulated unit			0.023
			(0.020)
Observations	9,671	9,671	9,671
Adjusted R^2	0.872	0.872	0.872

Note: State FE, Year FE, and Month FE are included. I add one to the inside of the logarithm because the covariates take 0 if there are no regulated units in a state in a particular time period. Standard errors are clustered at the state level.

A.3.1 Model

A generating unit can produce electricity with constant marginal cost mc_{jt} up to the capacity constraint of k_j. The marginal cost of unit j in time t is given by

$$mc_{jt}(\lambda) = p_{jt}^{fuel} \cdot HR_j + \lambda(1 - \alpha_{jt})R_{jt} \cdot HR_j.$$

Notations follow those in the main body of the paper. The first component of marginal cost $p_{jt}^{fuel} \cdot HR_j$ is fuel cost. The second component $\lambda(1-\alpha_{jt})R_{jt} \cdot HR_j$ is the cost associated with SO_2 emissions, which I call permit cost. The permit cost depends on the shadow value of permit λ. In the absence of transaction costs, the shadow value of emissions permits is equal to the market price of permits. In the structural model of the paper, the shadow value λ is endogenously determined due to the presence of transaction costs and permit banking.

I consider two approaches that endogenize the decision on electricity generation. First, I consider the cost minimization problem for each firm given their firm-level output. In period t, the firm decides the production allocation across generating units to minimize their total cost:

$$\min_{\{q_{jt}\}_{j=1}^{J_{it}}} \sum_{j=1}^{J_{it}} mc_{jt}(\lambda) \cdot q_{jt}$$
$$\text{s.t.} \sum_{j=1}^{J_{it}} q_{jt} = Q, q_{jt} \in [0, k_j]$$

Note that the total output at the firm level Q is exogenously given.

Without loss of generality, I order the generation units $i = 1, \cdots, N$ according to increasing marginal costs, i.e., $mc_{1t}(\lambda) \leq mc_{2t}(\lambda) \leq \cdots \leq mc_{Jt}(\lambda)$ in each period t. Then, the optimal choice of production allocation is given by

$$q_j = \begin{cases} k_j & \text{if } j = 1, \ldots, J^* - 1 \\ Q - \sum_{j=1}^{J^*-1} k_j & \text{if } j = J^* \\ 0 & \text{if } j = J^* + 1, \ldots \end{cases}$$

where J^* is the minimum number of generating units whose total generation capacity exceeds the given amount of total generation Q, i.e.,

$$J^* = \arg\min \left\{ J \,\middle|\, \sum_{j=1}^{J} k_j \geq Q \right\}.$$

Intuitively speaking, the firm operates the generation units with cheaper costs until it satisfies the total demand of Q.

I also construct a merit-order curve at the state level. Let $\mathcal{J}_{st}$ be the set of generating units located in state s. Then, given the state-level generation Q_{st}, the cost minimization problem is given by

$$\min_{\{q_{jt}\}_{j \in \mathcal{J}_{st}}} \sum_{j \in \mathcal{J}_{st}} mc_{jt}(\lambda) \cdot q_{jt}$$
$$\text{s.t.} \sum_{j \in \mathcal{J}_{st}} q_{jt} = Q_{st}, q_{jt} \in [0, k_j].$$

The solution to this problem is similarly given as that for the firm-level problem.

A.3.2 Descriptive Statistics of Fuel and Permit Cost

Before I analyze the endogenous response of electricity generation to cap-and-trade, I first report the descriptive statistics of the marginal cost in Table A3. To do this, I decompose the marginal cost into the fuel cost and the permit cost. To calculate the permit cost, I use the market price of emissions permits as shadow cost λ. I report the decomposition for three groups of generating units. First, the permit cost accounts for 7.9% and 6.2% of the marginal cost for Group I and Group II (coal). These numbers suggest that the introduction of cap-and-trade does not significantly increase the marginal cost of generation. Secondly, even if I consider the permit cost associated with SO_2 emissions, the marginal cost of coal units is substantially higher than that of gas and oil units.

Table A3: Summary Statistics of Fuel and Permit Cost

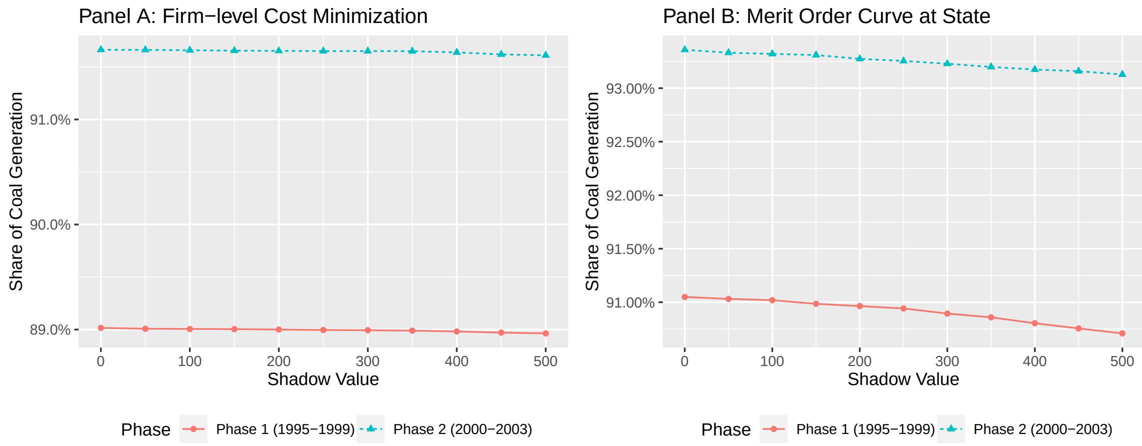
Variable	Mean	Std. Dev.	25 Percentile	Median	75 Percentile
			Group: Group I		
Fuel cost	16.29	15.55	11.41	13.16	15.97
Permit cost	1.39	1.19	0.52	1.08	1.93
			Group: Group II (Coal)		
Fuel cost	13.99	5.82	11.31	13.42	15.93
Permit cost	0.93	0.81	0.45	0.73	1.21
			Group: Group II (Gas and Oil)		
Fuel cost	54.58	48.37	35.59	46.65	59.71
Permit cost	0.17	0.41	0	0	0.04

Note: The unit is USD per MWh. The permit cost is the shadow cost of SO_2 emissions evaluated by the observed permit price.

A.3.3 Reallocation across generating units

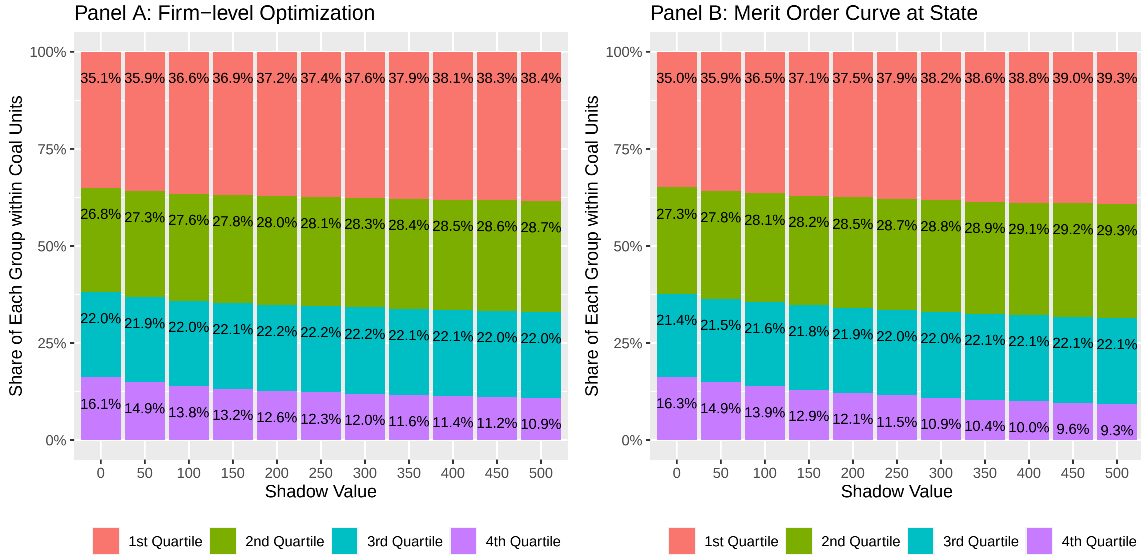
I now use the merit-order curve to investigate the endogenous change in the electricity generation. Figure A2 plots the share of coal generation as a function of the shadow value λ. In theory, the higher λ implies the higher costs of permits, leading to lower utilization of coal units. Although the share of coal generation is indeed decreasing as the shadow value λ increases, its magnitude is quite small. Even if the shadow value is set at \$500, which is quite high given that the permit price in my sample period is within the range of \$100–\$200, the share of coal generation only decreases by 0.2 percentage point in Panel B. The analysis implies that even though we consider the cost of emissions permits, reallocation of production across coal and other units (namely gas and oil units) is very limited. This finding is also consistent with the cost difference between coal and other units, as I discuss in Table A3.

Figure A2: Share of Coal Generation



I then investigate the reallocation of generations within coal units. Figure A3 illustrates how the generation share of each group of coal units changes according to the shadow value of emissions permits. Each group is defined by the quartiles of emissions rate of fuel (lbs/MMBtu). The lower quartile corresponds to cleaner units (i.e., lower emissions rate of fuel). The figure shows that reallocation from dirtier (i.e., 4th Quartile) to cleaner (i.e. 1st and 2nd Quartile) occurs when the shadow value of permits increases. Specifically, when the shadow value increases from 0 (i.e., no cap-and-trade) to 150 (which is in the ballpark of the observed permit prices in 2000–2003), the share of the dirtiest group falls from 16.1% to 13.2%, while that of the cleanest group rises from 35.1% to 36.9% in Panel A. This result indicates that introducing cap-and-trade may lead to the reallocation of generation within coal units to some extent. However, its magnitude is not sufficient to achieve the required level of emissions abatement under the Acid Rain Program.

Figure A3: Possible Reallocation within Coal Units in Phase II



Note: Quartile is defined by emissions rate of fuel (lbs/MMBtu) observed in the data. 1st quartile: lower than 0.56 lbs/MMBtu, 2nd quartile: higher than 0.56 lbs/MMBtu and lower than 1.00 lbs/MMBtu, 3rd quartile: higher than 1.00 lbs/MMbtu and lower than 1.70 lbs/MMbtu, 4th Quartile: higher than 1.70 lbs/MMBtu

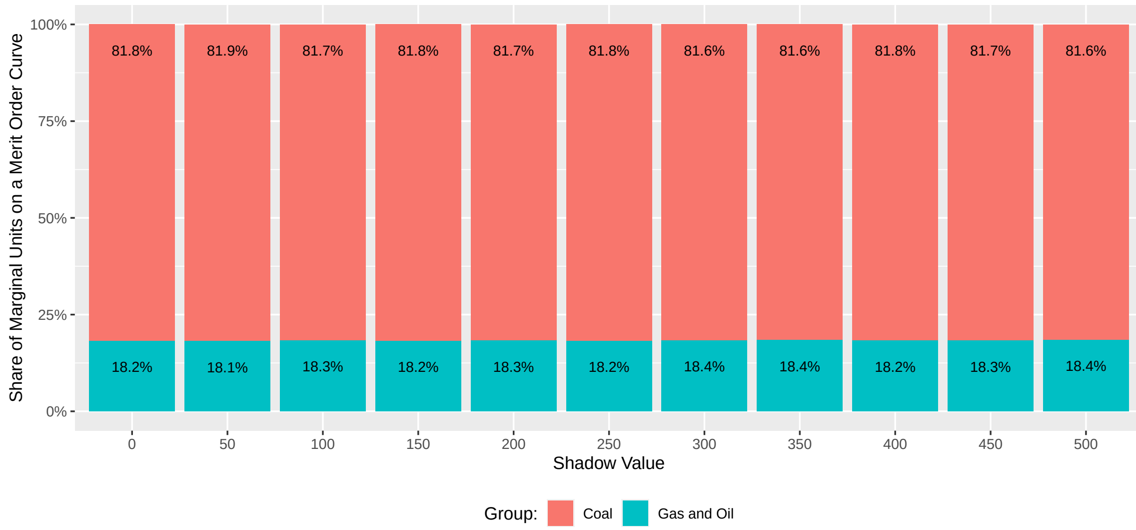
A.3.4 Marginal Units on a Merit-Order Curve at the State Level

Using a state-level merit-order curve for each month and year, I examine the share of marginal units. Figure A4 shows the share of marginal units as a function of the shadow value. Overall, the share of marginal coal units is around 82%, and this figure is almost constant across different values of the shadow values. Note that the share of coal units that are marginal is likely overestimated because I construct a merit-order curve using the monthly-level data, rather than high-frequency data (i.e., hourly-level data available in CEMS).[92]

This result raises the concern that the introduction of cap-and-trade could affect the output price determined by the marginal unit on a merit-order curve. Cap-and-trade increases the marginal cost of coal units due to the additional permit costs, which might be passed through to the output price. However, I believe this magnitude is likely to be relatively small. As I have shown in the descriptive statistics presented in Table A3, the permit cost is quite small compared to the fuel cost. Therefore, the potential impact of cap-and-trade on output price is likely to be quite small too.

[92]If I consider the demand fluctuation during the day, the model based on a merit-order curve produces a period of time when gas and oil units are marginal units.

Figure A4: Share of Marginal Units



B Online Appendix: Estimation of Fringe Demand

This subsection explains the estimation of the fringe demand function. I consider the following linear specification:

$$\bar{B}_t^{fringe} = \kappa_0 + \kappa_1 P_t + \kappa_2 Phase2_t, \tag{B.1}$$

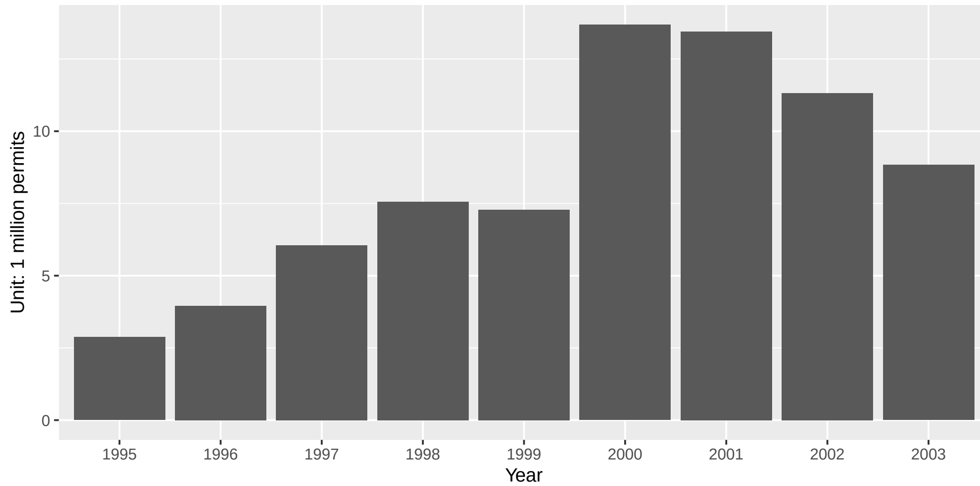
where $Phase2_t$ is the dummy for Phase II. The permit price P_t is subject to the endogeneity problem because the equilibrium permit price depends on the aggregate demand from the fringe firms. Thus, I use the sum of the initial allocation of permits owned by the firms in my sample as an instrument for P_t. The initial allocation of firms in the estimation sample is excluded from the fringe demand equation. Moreover, it is the part of the total amount of permits available in the market and thus affects permit prices. Table A4 reports IV estimates. Given the few data points available in my data (9 yearly observations from 1995 to 2003), the price coefficient is rather imprecisely estimated. For a comparison, Table A5 in the Online Appendix reports the first-stage result and the OLS result.

Table A4: Parameter Estimates of Fringe Demand

	Parameter	Description	Estimate	Standard Errors
Fringe Demand $\bar{B}_t(\cdot)$	κ_0	Constant	630110.45	577998.93
	κ_1	Permit Price	-3424.25	4133.94
	κ_2	Phase II dummy	-270591.92	191833.72

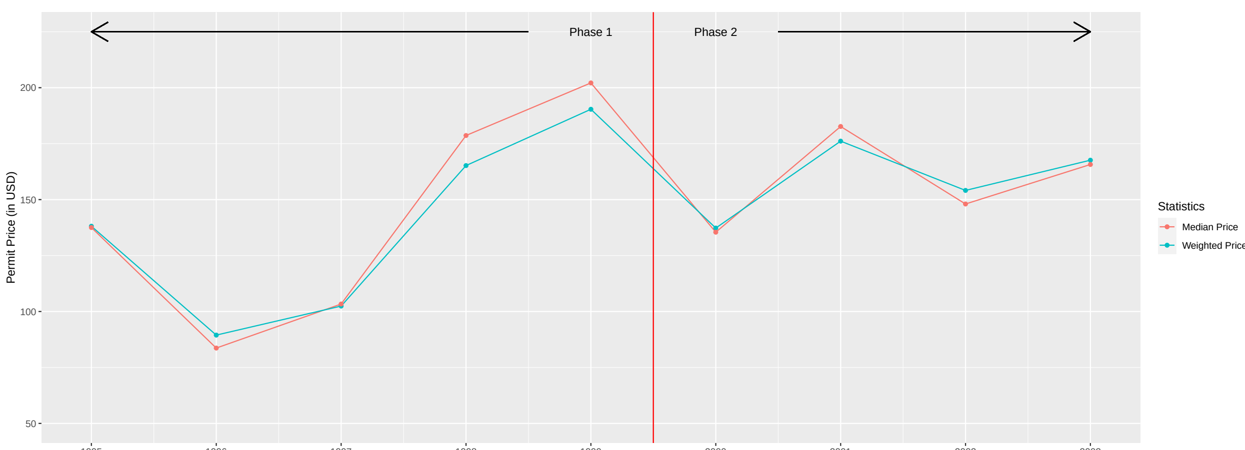
C Online Appendix: Additional Tables and Figures

Figure A5: Trading Volume over Time



Notes: The figure shows the aggregate volume of permit trading in each year. The unit is 1 million permits.

Figure A6: Comparison of volume-weighted mean and median prices of permits.



Notes: Prices are normalized to January 2000 prices using the producer price index. The prices are the weighted mean across months in each year. The weight is the aggregate trading volume of permits.

Figure A7: Frequency of Trading Normalized by Firm Size

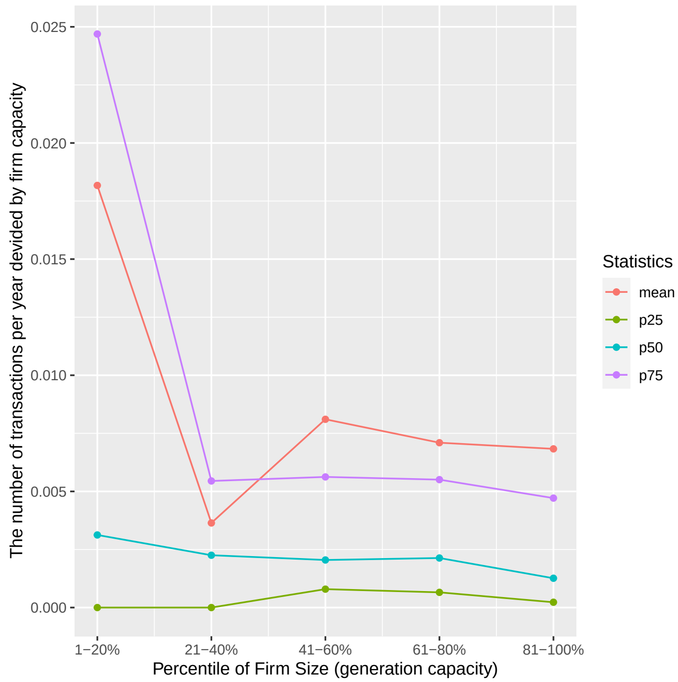
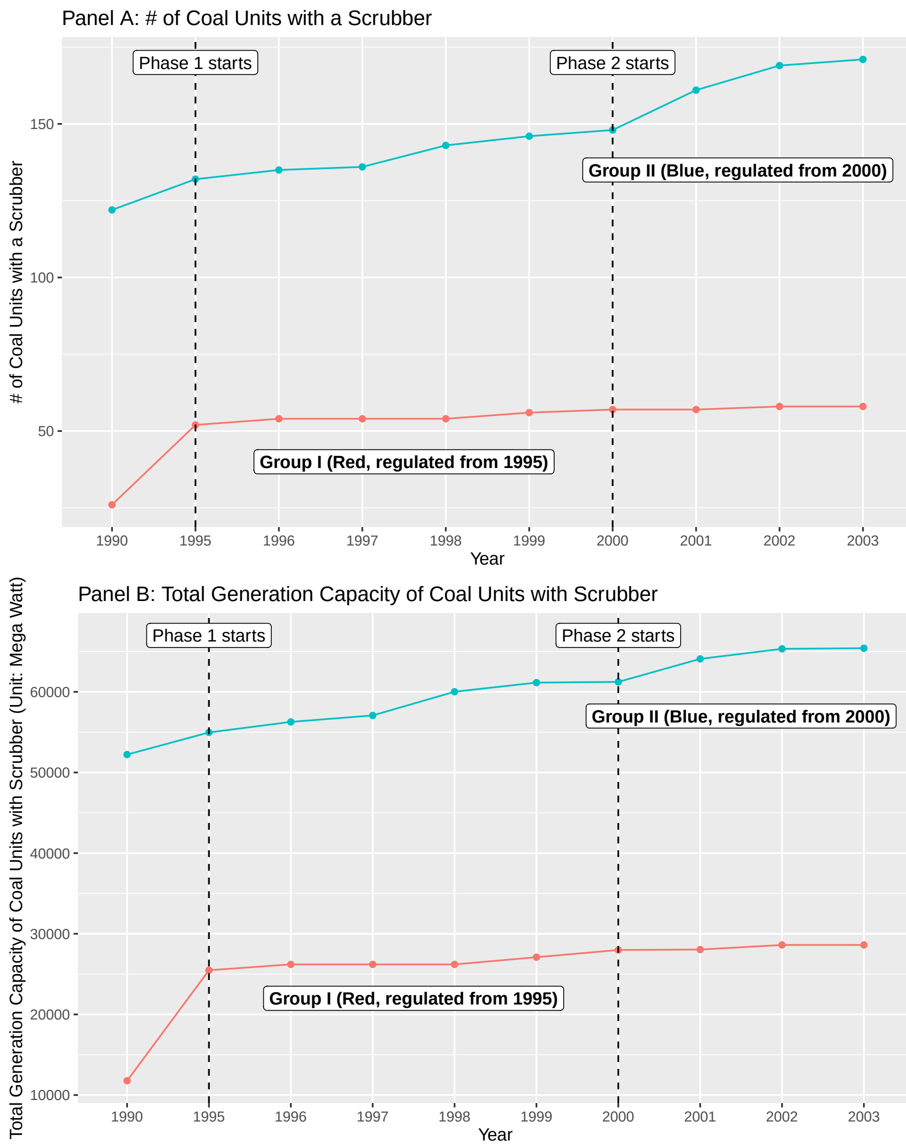


Figure A8: Scrubber Adoption



Notes: Panel A reports the number of coal units with a scrubber, while Panel B reports the total generation capacity of coal units with a scrubber. The Acid Rain Program was announced in 1990.

Table A5: Estimation Results of Fringe Demand

	OLS	1st Stage	IV
(Intercept)	974674.854	447.841	630110.447
	(377303.896)	(135.208)	(577998.931)
Permit Price	−5937.203		−3424.250
	(2642.120)		(4133.937)
Phase2 dummy	−216148.484	77.717	−270591.917
	(168181.241)	(30.112)	(191833.720)
Initial Allocation		−0.054	
		(0.024)	
R^2	0.620	0.531	0.563
Adj. R^2	0.494	0.375	0.417
Num. obs.	9	9	9

D Online Appendix: Details of Model Derivation

D.1 Derivation of the Optimality Conditions

In this appendix, I provide a detailed derivation of the optimality conditions for the constrained optimization problem introduced in Section 3.4. Recall that the constrained optimization problem is given by

$$\begin{aligned}
\max_{\{R_{jt}\}_{j=1}^{J_{it}}, b_{it}, h_{i,t+1}} \quad & \pi_{it}\left(\{q_{jt}, R_{jt}\}_{j=1}^{J_{it}}\right) - (P_t b_{it} + TC(|b_{it}|)) + \beta EV_{i,t+1}(h_{i,t+1}, 1, R_{i,t+1}) \\
\text{s.t.} \quad & e_{it}\left(\{q_{jt}, R_{jt}\}_{j=1}^{J_{it}}, \alpha_{it}\right) + h_{i,t+1} = a_{it} + h_{it} + b_{it}, \\
& h_{i,t+1} \geq 0.
\end{aligned}$$

The Lagrangian for this problem is

$$\begin{aligned}
\mathcal{L} =& \pi_{it}\left(\{q_{jt}, R_{jt}\}_{j=1}^{J_{it}}\right) - (P_t b_{it} + TC(|b_{it}|)) + \beta EV_{i,t+1}(h_{i,t+1}, 1, R_{i,t+1}) \\
&+ \lambda_{it}\left(a_{it} + h_{it} + b_{it} - e_{it}\left(\{q_{jt}, R_{jt}\}_{j=1}^{J_{it}}, \alpha_{it}\right) - h_{i,t+1}\right) + \mu_{it} h_{i,t+1},
\end{aligned}$$

where λ_{it} denotes the Lagrange multiplier on the transition of permit holding, $e_{it}\left(\{q_{jt}, R_{jt}\}_{j=1}^{J_{it}}, \alpha_{it}\right) + h_{i,t+1} = a_{it} + h_{it} + b_{it}$, and μ_{it} denotes the Lagrange multiplier on the nonborrowing constraint, $h_{i,t+1} \geq 0$. Taking the first-order conditions, I have

$$\frac{\partial \mathcal{L}}{\partial R_{jt}} = \frac{\partial \pi_{it}\left(\{q_{jt}, R_{jt}\}_{j=1}^{J_{it}}\right)}{\partial R_{jt}} - \lambda_{it} \frac{\partial e_{it}\left(\{q_{jt}, R_{jt}\}_{j=1}^{J_{it}}\right)}{\partial R_{jt}} = 0 \tag{D.1}$$

$$\frac{\partial \mathcal{L}}{\partial b_{it}} = -P_t - \frac{dTC(|b_{it}|)}{db_{it}} + \lambda_{it} = 0 \tag{D.2}$$

$$\frac{\partial \mathcal{L}}{\partial b_{it}} = \beta \frac{dEV_{i,t+1}(h_{i,t+1}, I_{i,t+1}, R_{i,t+1})}{dh_{i,t+1}} + \mu_{it} - \lambda_{it} = 0 \tag{D.3}$$

Equation (3.1) implies that $\frac{\partial \pi_{it}\left(\{q_{jt},R_{jt}\}_{j=1}^{J_{it}}\right)}{\partial R_{jt}} = -\frac{\partial p_{jt}^{fuel}(R_{jt})}{\partial R_{jt}} HR_j q_{jt}$ and Equation (3.2) implies that $\frac{\partial e_{it}\left(\{q_{jt},R_{jt}\}_{j=1}^{J_{it}}\right)}{\partial R_{jt}} = (1-\alpha_{jt})HR_j q_{jt}$. Thus, Equation (D.1) can be written as $\lambda_{it}\left((1-\alpha_{jt})HR_j q_{jt}\right) = -\frac{\partial p_{jt}^{fuel}(R_{jt})}{\partial R_{jt}} HR_j q_{jt}$, as shown in Equation (3.6). It is clear to see that Equations (D.2) and (D.3) can be written as Equations (3.7) and (3.8), respectively. Lastly, the complementary slackness condition with respect to the banking constraint $h_{it+1} \geq 0$ is given by $\mu_{it} h_{it+1} = 0, \mu_{it} \geq 0, h_{i,t+1} \geq 0$, as is shown in Equation (3.9).

D.2 Derivation of the Participation Probability $\mathbb{P}_{it}(h_{it}, \alpha_{it})$ and Ex-ante Value Function $EV_{it}(h_{it}, I_{it}, \alpha_{it})$

Participation probability of $\mathbb{P}_{it}(h_{it}, \alpha_{it})$ Recall that the participation probability is given by

$$\mathbb{P}_{it}(h_{it}, \alpha_{it}) = \int \mathbf{1}\left\{V_{it}^1(h_{it}, \alpha_{it}) - (F + \sigma_F \epsilon_{it}) > V_{it}^0(h_{it}, \alpha_{it})\right\} dG(\epsilon_{it}).$$

Since I assume the type-I extreme value distribution of ϵ_{it}, the participation probability is given by the well-known logit formula:

$$\mathbb{P}_{it}(h_{it}, \alpha_{it}) = \frac{\exp\left(\frac{V_{it}^1(h_{it},\alpha_{it})-F}{\sigma_F}\right)}{\exp\left(\frac{V_{it}^0(h_{it},\alpha_{it})}{\sigma_F}\right) + \exp\left(\frac{V_{it}^1(h_{it},\alpha_{it})-F}{\sigma_F}\right)}$$

Ex-ante value function of $EV_{it}(h_{it}, I_{it}, \alpha_{it})$ Recall that the ex-ante value functions are given by

$$EV_{it}(h_{it}, I_{it}, \alpha_{it}) = \begin{cases} \int \max\left\{V_{it}^0(h_{it}, \alpha_{it}), V_{it}^1(h_{it}, \alpha_{it}) - (F + \sigma_F \epsilon)\right\} dG(\epsilon) & if\ I_t = 0 \\ V_{it}^1(h_{it}, \alpha_{it}) & if\ I_t = 1. \end{cases}$$

Under the assumption that ϵ follows an i.i.d. type-I extreme value distribution, the expected value function when $I_{it} = 0$ can be written as

$$EV_{it}(h_{it}, I_{it} = 0, \alpha_{it}) = \sigma_F \log\left[\exp\left(\frac{V_{it}^0(h_{it}, \alpha_{it})}{\sigma_F}\right) + \exp\left(\frac{V_{it}^1(h_{it}, \alpha_{it}) - F}{\sigma_F}\right)\right].$$

By applying the Williams–Daly–Zachary theorem and the envelope theorem, the derivative of the expected value function with respect to the state variable h_{it} can be expressed as follows:

$$\frac{dEV_t(h_{it}, 0, \alpha_{it})}{dh_{it}} = \mathbb{P}_{it}(h_{it}, \alpha_{it})\lambda_{it}^1 + (1 - \mathbb{P}_t(h_{it}, \alpha_{it}))\lambda_{it}^0. \tag{D.4}$$

$$\frac{dEV_t(h_{it}, 1, \alpha_{it})}{dh_{it}} = \lambda_{it}^1, \tag{D.5}$$

where λ_{it}^1 and λ_{it}^0 are the Lagrange multipliers on the constraints for permit transitions in the optimization problems for traders and nontraders, respectively. I now provide a detailed derivation of the above equations.

Derivation of $\partial EV_t(h_t, I_t, \alpha_t)/\partial h_t$ I omit the index i for a firm for ease of exposition. I focus on the derivation of $\frac{\partial EV_t(h_t,0,\alpha_t)}{\partial h_t}$. Recall that

$$EV_t(h_t, 0, R_t) \quad = \quad \int \max\left\{V_t^0(h_t, R_t), V_t^1(h_t, R_t) - F - \sigma_F \epsilon\right\} dG(\epsilon).$$

By the chain rule, I have

$$\frac{dEV_t(h_t, 0, \alpha_t)}{dh_t} = \frac{\partial EV_t}{\partial V_t^0}\frac{dV_t^0}{dh_t} + \frac{\partial EV_t}{\partial V_t^1}\frac{dV_t^1}{dh_t}.$$

First, I derive $\frac{\partial EV_t}{\partial V_t^k}$ for $k = 0, 1$. This is an application of the Williams–Daly–Zachary theorem (see Theorem 3.1 in Rust, 1994). Using the interchange of integration and differentiation, I arrive at the following (I omit h_t for ease of exposition in the following derivation):

$$\begin{aligned}
\frac{\partial EV_t}{\partial V_t^1} &= \frac{\partial}{\partial V_t^1}\int \max\left\{V_t^1 - F - \sigma_F\epsilon, V_t^0\right\} dG(\epsilon) \\
&= \frac{\partial}{\partial V_t^1}\int_{\Upsilon^1}(V_t^1 - F - \sigma_F\epsilon)dG(\epsilon) + \frac{\partial}{\partial V_t^1}\int_{\Upsilon^0} V_t^0 dG(\epsilon) \\
&= \int_{\Upsilon^1}\frac{\partial}{\partial V_t^1}(V_t^{trade} - F - \sigma_F\epsilon)dG(\epsilon) + \int_{\Upsilon^0}\frac{\partial}{\partial V_t^1}V_t^0 dG(\epsilon) \\
&= \int_{\Upsilon^1} dG(\epsilon) \\
&= \mathbb{P}_t(\cdot),
\end{aligned}$$

where Υ^1 is the set of ϵ such that a firm chooses to participate (i.e., $\Upsilon^1 \equiv \{\epsilon : V_t^1 - F - \sigma_F\epsilon > V_t^0\}$), and Υ^0 is defined similarly. Note that I can apply a similar derivation to obtain $\frac{\partial EV_t}{\partial V_t^0} = 1 - \mathbb{P}(h_t)$.

Next, I calculate $\frac{\partial V_t^k}{\partial h_t}$, for $k = 0, 1$. The derivation is a direct application of the envelope theorem (or the Benveniste–Scheinkman formula):

$$\frac{\partial V_t^k}{\partial h_t} = \lambda_t^k,$$

where λ_{it}^k denotes the Lagrange multipliers in the corresponding optimization problems. Thus, I obtain

$$\frac{dEV_t(h_t, 0)}{dh_t} = \mathbb{P}_t(h_t)\lambda_t^1 + (1 - \mathbb{P}_t(h_t))\,\lambda_t^0.$$

D.3 Incentives in Abatement Investment

Here, I discuss how the incentive to invest in abatement is determined in the model. Using the envelope theorem, the marginal return from increasing the removal rate of a scrubber is given as follows:

$$\frac{\partial EV_{1995}}{\partial \alpha^1} = \sum_{t=1995}^{1999} \beta^{t-1995} \left(\lambda_{it} \cdot \sum_{j=1}^{J_{it}} R_{jt} \cdot HR_{jt} \cdot q_{jt} \right) - \beta^{2000-1995} \frac{\partial}{\partial \alpha^1} \Gamma(\alpha^2 - \alpha^1).$$

The first term is the returns from emissions abatement evaluated at the shadow value λ_{it}. The second term is the saving of investment costs in Phase II due to the earlier investment in Phase I.

The primary component in the return on investment is the first term. By increasing the removal rate of a scrubber, a firm can marginally reduce its emissions by $\sum_{j=1}^{J_{it}} R_{jt} \cdot HR_{jt} \cdot q_{jt}$. This marginal abatement is evaluated at the shadow value of λ_{it}. Thus, the return on investment is given by the discounted sum of the returns on the marginal abatement. The path of shadow values λ_{it} is key for the investment incentives. As discussed in Section 3.7, the shadow value λ_{it} and equilibrium permit price P_t are affected by both permit banking and transaction costs.

D.4 Details of Decomposition of Change in Health and Environmental Damages

The aggregate health and environmental damage is given by

$$D = \sum_{t=1995}^{2003} \beta^{t-1995} \left(\sum_{i=1}^{N} \sum_{j=1}^{J_{it}} d_j e_{jt} \right),$$

where e_{jt} is emissions from unit j in year t and d_j is the health and environmental damage from emissions produced by unit j. Note that d_j is the county-level estimates of SO_2 damage constructed by Muller and Mendelsohn (2009).

I rewrite the aggregate damage as

$$\begin{aligned} D &= \sum_{t=1995}^{2003} \beta^{t-1995} E_t \left(\frac{\sum_{i=1}^{N} \sum_{j=1}^{J_{it}} d_j e_{jt}}{E_t} \right) \\ &= \sum_{t=1995}^{2003} \tilde{d}_t E_t, \end{aligned}$$

where $E_t = \sum_{i=1}^{N} \sum_{j=1}^{J_{it}} e_{jt}$ is the aggregate emissions in year t and $\tilde{d}_t = \beta^{t-1995} \left(\frac{\sum_{i=1}^{N} \sum_{j=1}^{J_{it}} d_j e_{jt}}{E_t} \right)$ is the (discounted) average SO_2 damages in year t. Note that $\tilde{d}_t$ interpreted as the weighted average of health and environmental damages, where the weight is given by the amount of SO_2 emissions.

E Online Appendix: Computational Details of Solving the Model

Appendix E explains the computational procedure used to solve the structural model.

E.1 Decomposition of the Per-Period Problem

One of the choice variables in the individual dynamic decision problem is the unit-level coal quality R_{jt}, which appears in the profit function π_{it}, given by Equation (3.1), and the firm-level emissions, $e_{it} = \sum_{j=1}^{J_{it}}(1-\alpha_{jt})R_{jt} \cdot HR_j \cdot q_{jt}$. Because each firm has multiple generation units, solving unit-level production in a dynamic framework seems computationally demanding. Therefore, to reduce the computational burden, I decompose the per-period problem into the following two problems. First, I consider the following optimization problem with respect to the unit-level coal quality $\{R_{jt}\}_{j \in J_{it}}$, *holding firm-level emissions e_{it} fixed:*

$$\begin{aligned} \Pi_{it}(e_{it}, \alpha_{it}) \equiv \quad & \max_{\{R_{jt}\}_{j \in J_{it}}} \quad \pi_{it}\left(\{q_{jt}, R_{jt}\}_{j=1}^{J_{it}}\right) \\ & \text{s.t.} \qquad \sum_{j=1}^{J_{it}}(1-\alpha_{it})R_{jt} \cdot HR_j \cdot q_{jt} = e_{it}. \end{aligned}$$

$\Pi_{it}(e_{it}, \alpha_{it})$ is the optimal profit *as a function of the firm-level emissions e_{it}.* Note that the FOCs for this subproblem are

$$\begin{aligned} \lambda_{it}^{sub}\left((1-\alpha_{jt})HR_j q_{jt}\right) &= -\frac{\partial p_{jt}^{fuel}(R_{jt})}{\partial R_{jt}} HR_j q_{jt} \\ \sum_{j=1}^{J_{it}}(1-\alpha_{it})R_{jt} \cdot HR_j \cdot q_{jt} &= e_{it}, \end{aligned}$$

where λ_{it}^{sub} is the Lagrange multiplier of the constraint on firm-level emissions in the above problem.

I now use $\Pi_{it}(e_{it}, \alpha_{it})$ to consider the dynamic decision problem:

$$\begin{aligned} \max_{e_{it}, b_{it}, h_{i,t+1}} \quad & \Pi_{it}(e_{it}, \alpha_{it}) - (P_t b_{it} + TC(b_{it})) + \beta EV_{i,t+1}(h_{i,t+1}, 1, R_{i,t+1}) \\ \text{s.t.} \quad & e_{it} + h_{i,t+1} = a_{it} + h_{it} + b_{it}, \\ & h_{i,t+1} \geq 0. \end{aligned}$$

Note that the choice variables are now reduced to e_{it}, b_{it}, and $h_{i,t+1}$.

When I numerically solve the individual dynamic decision problem, I follow two steps. First, I construct $\Pi_{it}(e_{it}, \alpha_{it})$ using the unit-level FOC for production. I then use the precomputed $\Pi_{it}(e_{it}, \alpha_{it})$ to solve the individual dynamic decision problems.

E.2 Individual Optimization

I explain the computational procedure for solving an individual problem. For notational simplicity, I omit the script i for a particular firm. Because the model has a finite period, it can be solved using backward induction.

1. Phase II (2003 to 2000): I solve the optimization problem from 2003 to 2000. Note that I use $CV_{T+1}(h_{T+1}, \alpha^2)$ as a continuation value in the terminal period 2003. By solving with backward induction, I obtain the policy function $\hat{x}_t(h_t, I_t, \alpha^2)$ for emissions e_t, net purchase b_t, and banking h_{t+1}, and the expected value function in 2000 $EV_{2000}(h_{2000}, I_{2000}, \alpha^2)$.

2. Investment decision for Phase II: I define the continuation value at the timing of making the investment decision for Phase II by $W_{2000}(h_{2000}, I_{2000}, \alpha^1)$. The decision problem is given by

$$
\begin{aligned}
W_{2000}(h_{2000}, I_{2000}, \alpha^1) \equiv \quad & \max_{\alpha^2} \quad EV_{2000}(h_{2000}, I_{2000}, \alpha^2) - \Gamma(\alpha^2, \alpha^1). \\
& \text{s.t.} \quad \alpha^2 \leq \alpha^1
\end{aligned}
$$

   By solving this problem, I obtain the investment policy function $\alpha^{2*}(h_{2000}, I_{2000}, \alpha^1)$.

3. Phase I (1999 to 1995): I repeat the same procedure as that in step 1. Note that the continuation value in the problem at $t = 1999$ is given by $W_{2000}(h_{2000}, I_{2000}, \alpha^1)$.

4. Investment for Phase I: The problem is given by

$$
\begin{aligned}
& \max_{\alpha^1} EV_{1995}(0, 0, \alpha^1) - \Gamma(\alpha^1, \alpha^0). \\
& \text{s.t.} \alpha^1 \leq \alpha^0
\end{aligned}
$$

   Note that $h_{1995} = 0$ and $I_{1995} = 0$ in 1995.

E.3 Computation of a Dynamic Competitive Equilibrium

The computational procedure for finding an equilibrium is parallel to the estimation procedure introduced in Section 4.

1. Fix a candidate of permit prices: $\mathbf{P} = \{P_t\}_{t=1995}^{2003}$.

2. Solve the individual problem using backward induction and obtain the policy function $\hat{x}_{it}(h_{it}, I_{it}, \alpha_{it})$ for emissions e_t, net purchase b_t, and banking h_{t+1}, participation probability $P_{it}(h_{it}, \alpha_{it})$, and the investment decisions $\alpha_i^1(h_{i1995}, I_{i1995})$ and , $\alpha_i^2(h_{i,2000}, I_{i,2000}, \alpha_i^1)$.

3. Consider the timing of market participation. Denote the year of participation by $s \in \{\emptyset, 1995, \cdots, 2003\}$. Here, $s = \emptyset$ means that a firm does not trade in a period.

4. For each path of participation timing, I simulate the optimal decisions using the policy functions.

5. Calculate the probability that each path of participation timing is realized.

6. The simulated optimal decisions are given as

$$\hat{x}_{it} = \sum_{s \in \{\emptyset, 1995, \cdots, 2003\}} Prob_i^{enter}(s)\hat{x}_{it}(s),$$

where x denotes the choice variables.

7. Check the market-clearing condition as

$$\sum_{i=1}^{N} \hat{b}_{it}(\mathbf{P}) + \bar{B}_t^{fringe}(P_t) = 0 \ \forall t = 1995, \cdots, 2003.$$

8. Stop the iteration when the following condition is satisfied:

$$\max_{t=1995,\cdots,2003} \left| \sum_i \hat{b}_{it}(\mathbf{P}) + \bar{B}_t^{fringe}(P_t) \right| < 1000.$$

Note that this criterion is sufficiently tight to ensure that the absolute value of the price change is in the order of magnitude of 1e-1.

9. If the above is not satisfied, repeat steps 1–7 with the updated price vector (explained below), until the market-clearing conditions are satisfied.

Price Update Rule To update the price in each iteration, I construct the following heuristic rule that exploits the market-clearing conditions and the optimality conditions. Denote the current candidate of an equilibrium price vector by $\mathbf{P}^l = \{P_t^l\}_{t=1995}^{2003}$. The next candidate of price in year t, P_t^{l+1}, is given by solving the following equation:

$$\sum_{i=1}^{N} \sum_{s \in \{\emptyset, 1995, \cdots, 2003\}} P_{i,enter}(s) \cdot TC'^{(-1)} \left(\hat{\lambda}_{it}(\mathbf{P}^l, s) - P_t^{l+1} \right) + \bar{B}_t^{fringe}(P_t^{l+1}) = 0,$$

where $\hat{\lambda}_{it}(\mathbf{P}^l, s)$ is the prediction of the shadow value when the current price candidate is $\mathbf{P}^l$ and the year of participation is s. Note that at the fixed point of this equation, where $\mathbf{P}^l = \mathbf{P}^{l+1}$,

$$TC'^{(-1)} \left(\hat{\lambda}_{it}(\mathbf{P}^l, s) - P_t^l \right) = b_{it}(\mathbf{P}^l, s),$$

such that the market-clearing conditions are satisfied in all periods.

The computation procedure with this price update rule works relatively well in numerical simulations. The algorithm finds an equilibrium price vector in fewer than 10 iterations in most cases, though I do not have a formal proof of this property of the algorithm.

F Online Appendix: Details of Counterfactual Simulations

F.1 Eliminating Transaction Costs

I now consider the case with permit banking. In the absence of transaction costs, Rubin (1996) has shown that the equilibrium path of permit prices grows at the rate of β^{-1}, as long as the aggregate banking is positive, which implies that

$$\begin{aligned} P_{t+1} &= \beta^{-1} P_t \\ \Longleftrightarrow P_t &= \beta^{-(t-1)} P_{1995} \text{ for } t \in \{1995, \cdots, 2003\}. \end{aligned}$$

The optimal decision on emissions, given the emissions rate of fuel, is determined by $\partial \pi_{it} / \partial R_{jt} = P_t \; \forall j$. As discussed in Section 3.7.2, individual decisions on net purchases and banking are not determined from the model because the current shadow value $\lambda_t = P_t$ is equal to the discounted marginal value of banking $\beta \lambda_{t+1} = \beta P_{t+1} = P_t$. In other words, banking and trading decisions are arbitrary as long as a firm can produce the level of emissions determined by the optimality condition.

Now, I consider the investment decisions. The continuation value at the beginning of Phase II is given by

$$\begin{aligned} V_{i,2000}(h_{i,2000}, \alpha_i^2) &= \sum_{t=2000}^{2003} \beta^{t-2000} \left[\pi_{it} \left(\{q_{jt}\}_{j=1}^{J_{it}}, \alpha_i^2 \right) - P_t b_{it} \right] + \beta^{2003-2000} CV(h_{i,T+1}) \\ &= \sum_{t=2000}^{2003} \beta^{t-2000} \left[\pi_{it} \left(\{q_{jt}\}_{j=1}^{J_{it}}, \alpha_i^2 \right) - P_t \cdot (e_{it} - a_{it}) \right] \\ &\quad + \beta^{2003-2000} \left\{ CV(h_{i,T+1}) - P_T h_{i,T+1} \right\} \\ &\quad + \sum_{t=2000}^{2003} \beta^{t-2000} P_t h_{it} + \sum_{t=2000}^{2002} \beta^{t-2000} P_t h_{it+1} \\ &= \sum_{t=2000}^{2003} \beta^{t-2000} \left[\pi_{it} \left(\{q_{jt}\}_{j=1}^{J_{it}}, \alpha_i^2 \right) - P_t \cdot (e_{it} - a_{it}) \right] \\ &\quad + \beta^{2003-2000} \left\{ CV(h_{i,T+1}) - P_T h_{i,T+1} \right\} + P_{2000} h_{i,2000}, \end{aligned}$$

where the last equality uses the equilibrium relationship $\beta P_{t+1} = P_t$. The investment problem is

$$
\begin{aligned}
W_{i,2000}(h_{i,2000}, \alpha_i^1) = \quad & \max_{\alpha_i^2} \quad V_{2000}(h_{i,2000}, \alpha_i^2) - \Gamma(\alpha_i^2, \alpha_i^1). \\
& s.t. \quad \alpha_i^2 \leq \alpha_i^1.
\end{aligned}
$$

Note that $h_{i,2000}$ does not affect the optimal investment level of α_i^2.

The continuation value at the beginning of Phase I is given as

$$
\begin{aligned}
V_{1995}(h_{i,1995}, \alpha_i^1) \quad = \quad & \sum_{t=1995}^{1999} \beta^{t-1995} \left[\pi_{it} \left(\{q_{jt}\}_{j=1}^{J_{it}}, \alpha_i^1 \right) - P_t(e_{it} - a_{it}) \right] \\
& + \beta^{1999-1995} \left(\beta W_{2000}(h_{i,2000}, \alpha_i^1) - P_{1999} h_{i,2000} \right).
\end{aligned}
$$

The investment problem is similar to that in Phase II.

Finally, I consider the market-clearing condition. By aggregating the transition equation of permit holding (3.3) over individual firms and time, I have

$$
\sum_{t=1995}^{2003} E_t(P_t) + H_{T+1} = \sum_{t=1995}^{2003} A_t + \sum_{t=1995}^{2003} B_t, \tag{F.1}
$$

where $E_t = \sum_{i=1}^{N} e_{it}(P_t)$, and other uppercase variables are defined similarly. The market-clearing condition in each period is

$$
B_t + \bar{B}_t^{fringe}(P_t) = 0.
$$

By substituting this condition into Equation (F.1), I have

$$
\sum_{t=1995}^{2003} E_t \left(\beta^{-(t-1)} P_{1995} \right) + H_{T+1}(\beta^{-(T-1)} P_{1995}) = \sum_{t=1995}^{2003} A_t + \sum_{t=1995}^{2003} -\bar{B}_t^{fringe} \left(\beta^{-(t-1)} P_{1995} \right).
$$

The equilibrium price P_{1995} is determined by this equation and, thus, so is the whole path of the equilibrium price.

F.2 Model without Permit Banking between Phase I and II

I explain the case in which firms are not allowed to bank emissions permits between Phases I and II. The decision problem is the same as that introduced in Section 3, except for 1999, the last year of Phase I.

I first consider the problem for a trader in 1999 (i.e.,$t = 1999$). I omit the subscript i for simplicity. The problem is given by

$$V^1_{1999}(h_{1999}, I_{1999}=1, \alpha^1) = \max_{\{R_{jt}\}_{j=1}^{J_{it}}, b_t} \quad \pi_t\left(\{q_{jt}, R_{jt}\}_{j=1}^{J_{it}}\right) - (P_t b_t + TC(|b_t|)) + \beta W_{2000}(0, I_{2000}, \alpha^1)$$
$$\text{s.t.} \quad e_t\left(\{q_{jt}, R_{jt}\}_{j=1}^{J_{it}}, \alpha^1\right) = a_t + h_t + b_t.$$

Note that permit banking h_{2000} is not among the choice variables, while the continuation value $W_{2000}(0, I_{2000}, \alpha^1)$ is evaluated at $h_{2000} = 0$. The optimality conditions of the problem are given by Equations (3.6) and (3.7).

Next, consider the case in which a firm is a non-trader:

$$V^0_{1999}(h_{1999}, I_t=0, \alpha^1) = \max_{\{R_{jt}\}_{j=1}^{J_{it}}, b_t} \quad \pi_t\left(\{q_{jt}, R_{jt}\}_{j=1}^{J_{it}}\right) + \beta W_{2000}(0, I_{2000}, \alpha^1)$$
$$\text{s.t.} \quad e_t\left(\{q_{jt}, R_{jt}\}_{j=1}^{J_{it}}, \alpha^1\right) \leq a_t.$$

In this case, a firm may not consume all its permits owing to the capacity constraints of production. The emissions level is given by

$$e^*_t = \min\{a_t, e^{max}_t\},$$

where e^{max}_t is the emissions level when a firm faces zero shadow costs of permits $\lambda_t = 0$.

Other components, including the participation and the investment decisions, are the same as in the baseline case (i.e., the case that includes both permit banking and transaction costs).

Reference for Online Appendix

Asker, John, Allan Collard-Wexler, and Jan De Loecker (2019). "(Mis) allocation, market power, and global oil extraction". In: *American Economic Review* 109.4, pp. 1568–1615.

Borenstein, Severin, James B Bushnell, and Frank A Wolak (2002). "Measuring market inefficiencies in California's restructured wholesale electricity market". In: *The American Economic Review* 92.5, pp. 1376–1405.

Bushnell, James B, Erin T Mansur, and Celeste Saravia (2008). "Vertical arrangements, market structure, and competition: An analysis of restructured US electricity markets". In: *The American Economic Review* 98.1, pp. 237–266.

Callaway, Brantly and Pedro HC SantAnna (2021). "Difference-in-differences with multiple time periods". In: *Journal of Econometrics* 225.2, pp. 200–230.

Energy Information Administration (1990–2003). *Form EIA-826 detailed data.* (last accessed: September 17, 2014). URL: `http://www.eia.gov/electricity/data/eia826/`.

Goodman-Bacon, Andrew (2021). "Difference-in-differences with variation in treatment timing". In: *Journal of Econometrics* 225.2, pp. 254–277.

Muller, Nicholas Z and Robert Mendelsohn (2009). "Efficient pollution regulation: getting the prices right". In: *The American Economic Review* 99.5, pp. 1714–1739.

Rubin, Jonathan D (1996). "A model of intertemporal emission trading, banking, and borrowing". In: *Journal of Environmental Economics and Management* 31.3, pp. 269–286.

Rust, John (1994). "Structural estimation of Markov decision processes". In: *Handbook of econometrics* 4, pp. 3081–3143.

Sun, Liyang and Sarah Abraham (2021). "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects". In: *Journal of Econometrics* 225.2, pp. 175–199.