# Appendices

## A Proofs of Propositions

To simply exposition, I assume without loss of generality that the the discount rate $r = 0$. The firms' problem in the first period can then be written as:

$$\min_{a_1, I_1} \quad P_1 \cdot (\bar{e} - a_1) + C(a_1, I_1) + K \cdot I_1 + \mathbb{E}\Big[\min_{a_2, I_2}\{P_2 \cdot (\bar{e} - a_2) + C(a_2, I_2) + K \cdot (I_2 - I_1)\}\Big]$$
$$s.t. \quad a_t \in [0, \bar{e}], \quad I_t \in \{0, 1\}, \quad I_2 \geq I_1$$

### A.1 Proof of Proposition 1

Firms are differentiated by their costs of capital. Specifically, there are three groups of firms: (1) firms that will install the capital technology in the first period, (2) firms that will wait to install the capital technology only if the high emissions price is realized, (3) firms that will never install the capital technology. The third group will have the highest cost of capital and regardless of $\rho$ they will never install the technology. Therefore, the share of firm's adopting the technology in the first period will be determined by the cutoff capital cost $K_1^*$ that separates the first and second groups. I will show that $K_1^*$ is increasing in $\rho$ and therefore the the share of adopters in period one $F(K_1^*)$ will also be increasing in $\rho$.

The expected net benefits from installing in period 1 are equal to: $P_1 \cdot a_1^I - C(a_1^I, 1) - K + E[\min_{a_2, I_2}\{P_2 \cdot a_2 - C(a_2, I_2)\} | I_1 = 1]$.[37] If the firm installs the technology in period 1, it saves on permit costs and abatement costs in period 1 and anticipates saving on permit costs and abatement costs in period 2; however, it also must pay the capital cost $K$ and the abatement cost $C(a_1^I, 1)$. If the firm waits until period 2 and only installs the capital technology if the stringent price is enacted, then its expected benefit will be $P_1 \cdot a_1^N - C(a_1^N, 0) + E[\min_{a_2, I_2}\{P_2 \cdot a_2 - C(a_2, I_2) + K \cdot I_2\} | I_1 = 0]$. Firms should invest if:

$$\begin{aligned} & P_1 \cdot a_1^I - C(a_1^I, 1) - K + \mathbb{E}[\min_{a_2, I_2}\{P_2 \cdot a_2 - C(a_2, I_2)\} | I_1 = 1] \\ & \geq P_1 \cdot a_1^N - C(a_1^N, 0) + \mathbb{E}[\min_{a_2, I_2}\{P_2 \cdot a_2 - C(a_2, I_2) + K \cdot I_2\} | I_1 = 0] \end{aligned} \tag{9}$$

Let $a_2^{IH}$ be the optimal level of abatement in period 2, conditional on having installed the capital technology ($I_2 = 1$) and emission prices being high. Let $a_2^{IL}$ be the optimal abatement level if permit prices are low and the firm has installed the technology. Furthermore, let $a_2^{NH}$ and $a_2^{NL}$ be the optimal abatement levels for firms who have not installed the technology for the high and low emission price cases respectively. Expanding the expectations, we have:

[37] Recall $a_1^I$ is the optimal first period abatement conditional on installing the capital technology, and $a_1^N$ is the optimal first period abatement choice for firms that do not install the technology.

$$
\begin{aligned}
&P_1 \cdot a_1^I - C(a_1^I, 1) - K + \rho(P_2^H \cdot a_2^{IH} - C(a_2^{IH}, 1)) \\
&+(1-\rho)(P_2^L \cdot a_2^{IL} - C(a_2^{IL}, 1)) \geq P_1 \cdot a_1^N - C(a_1^N, 0) \\
&+\rho(P_2^H \cdot a_2^{IH} - C(a_2^{IH}, 1) - K) + (1-\rho)(P_2^L \cdot a_2^{NL} - C(a_2^{NL}, 0))
\end{aligned} \tag{10}
$$

To understand the the right-hand side of the inequality, recall that firms in the second group will install the capital technology in the second period only if the high permit price is realized, which will occur with probability $\rho$. The cutoff cost $K_1^*$ for investment in the first period is determined by the capital cost at which a firm would be indifferent between investing and waiting. Setting the right and left hand sides of (10) equal to each other and solving for $K$ we obtain:

$$
\begin{aligned}
K_1^* = &\frac{\Big(P_1 \cdot a_1^I - C(a_1^I, 1)\Big) - \Big(P_1 \cdot a_1^N - C(a_1^N, 0)\Big)}{1-\rho} \\
&+\Big(P_2^L \cdot a_2^{IL} - C(a_2^{IL}, 1)\Big) - \Big(P_2^L \cdot a_2^{NL} - C(a_2^{NL}, 0)\Big)
\end{aligned} \tag{11}
$$

differentiating (11) with respect to $\rho$ we have:

$$
\begin{aligned}
\frac{dK_1^*}{d\rho} = &\frac{P_1}{1-\rho}\Big(\frac{da_1^I}{d\rho} - \frac{da_1^N}{d\rho}\Big) + \frac{1}{1-\rho}\Big(C_a(a_1^N, 0)\frac{da_1^N}{d\rho} - C_a(a_1^I, 1)\frac{da_1^I}{d\rho}\Big) \\
&+\frac{1}{(1-\rho)^2}\Big([P_1 \cdot a_1^I - C(a_1^I, 1)] - [P_1 \cdot a_1^N - C(a_1^N, 0)]\Big) + P_2^L\Big(\frac{da_2^{IL}}{d\rho} - \frac{da_2^{NL}}{d\rho}\Big) \\
&+\Big(C_a(a_2^{NL}, 0)\frac{da_2^{NL}}{d\rho} - C_a(a_2^{IL}, 1)\frac{da_2^{IL}}{d\rho}\Big)
\end{aligned} \tag{12}
$$

Substituting in the equilibrium conditions, $C_a(a_1, I_1) = P_1$ and $C_a(a_2, I_2) = P_2$, and canceling terms we are left with:

$$
\frac{dK_1^*}{d\rho} = \frac{1}{(1-\rho)^2}\Big([P_1 \cdot a_1^I - C(a_1^I, 1)] - [P_1 \cdot a_1^N - C(a_1^N, 0)]\Big) \tag{13}
$$

We know that the first term in brackets must be larger than the second term in brackets. To see this, notice $P_1 \cdot a_1^I - C(a_1^I, 1) \geq P_1 \cdot a_1^N - C(a_1^N, 1)$ since $a_1^I$ is the optimal abatement choice conditional on having $I_1 = 1$ by definition. Additionally we know that $P_1 \cdot a_1^N - C(a_1^N, 1) > P_1 \cdot a_1^N - C(a_1^N, 0)$ which follows from the assumption that marginal cost of abatement is lower once the capital technology is installed. This means the term in the large parentheses is positive, and since $\frac{1}{(1-\rho)^2}$ is positive this implies $\frac{dK_1^*}{d\rho} > 0$. Finally, since the cumulative distribution function $F$ must be non-decreasing in its argument it follows that $\frac{dF(K_1^*)}{d\rho} \geq 0$

■

## A.2 Proof of Proposition 2

We will show that $\frac{d\mathbf{e_1}}{d\rho} < 0$. Total emissions is equal to the sum of emissions from firms that invest in the technology and emissions from those that do not:

$$\mathbf{e_1} = M\Big(F(K_1^*)(\overline{e} - a_1^I) + (1 - F(K_1^*))(\overline{e} - a_1^N)\Big) \tag{14}$$

We next differentiate with respect to $\rho$ to obtain a comparative static:

$$\frac{d\mathbf{e_1}}{d\rho} = M\Big[f(K^*)\frac{dK^*}{d\rho}((\overline{e} - a_1^I) - (\overline{e} - a_1^N)) - F(K^*)\frac{da_1^I}{d\rho} - (1 - F(K^*))\frac{da_1^N}{d\rho}\Big] \tag{15}$$

where $f$ is the probability density function of $K$. We know that the first term in the brackets, $f(K^*)\frac{dK^*}{d\rho}((\overline{e} - a_1^I) - (\overline{e} - a_1^N))$, is negative since $f$ is non-negative by definition, $\frac{dF(K_1^*)}{d\rho}$ is positive as shown above, and $((\overline{e} - a_1^I) - (\overline{e} - a_1^N))$ is negative because firms that install the technology will have lower emissions. The next two terms in the brackets are equal to zero because $\frac{da_1^I}{d\rho} = 0$ and $\frac{da_1^N}{d\rho} = 0$, this can be shown by differentiating the first order condition $C_a(a_1, I_1) = P_1$ with respect to $\rho$. Therefore, since $M$ is also non-negative, it must be the case that $\frac{d\mathbf{e_1}}{d\rho} \leq 0$.

■

## A.3 Proof of Proposition 3

Let $\mathbf{a_1^N}$ denote total abatement by firms that do not adopt the technology in the first period, $\mathbf{a_1^N} = \sum_i a_{i1} \cdot \mathbb{1}(I_{i1} = 0)$. Total abatement by non-adopters is equal to:

$$\mathbf{a}_1^N = M\Big((1 - F(K_1^*))(a_1^N)\Big) \tag{16}$$

Differentiating with respect to $\rho$ we have:

$$\frac{d\mathbf{a}_1^N}{d\rho} = M\Big[\underbrace{-f(K_1^*)\tfrac{dK_1^*}{d\rho}(a_1^N)}_{[1]} + \underbrace{(1 - F(K_1^*))\tfrac{da_1^N}{d\rho}}_{[2]}\Big] \tag{17}$$

As the probability of the high price regime increases, more firms adopt the technology, which works to reduce total abatement by non-adopters, this effect is labeled [1] in equation 17. This term is negative since $f(K_1^*)$, $a_1^N$ are positive and $\frac{dK_1^*}{d\rho}$ is positive by proposition 1. Since $\frac{da_1^N}{d\rho} = 0$, the term labeled [2] in equation 17 equals zero. Therefore, $\frac{d\mathbf{a}_1^N}{d\rho} \leq 0$.

## A.4 Proof of Proposition 4

Define $K_2^*$ as the the cutoff capital cost that a firm would be indifferent to installing the technology in the second period, conditional on the high price regime occurring:

$$K_2^* = (P_2^H \cdot a_2^{IH} - C(a_2^{IH}, 1)) - (P_2^H \cdot a_2^{NH} - C(a_2^{NH}, 0)) \tag{18}$$

Notice the cutoff does not depend on $\rho$ since the uncertainty has already been resolved at this point. The number of firms that adopt is the number of firms that have capital costs smaller than $K_2^*$ but have capital costs larger than $K_1^*$ (i.e., did not invest in the first period). This number of firms can be expressed as:

$$M * \max\{0, (F(K_2^*) - F(K_1^*))\} \tag{19}$$

Let $\hat{\rho}$ be defined such that $K_2^* = K_1^*(\hat{\rho})$. Then differentiating (19) we obtain:

$$\frac{d[M * \max\{0, (F(K_2^*) - F(K_1^*(\rho)))\}]}{d\rho} = \begin{cases} 0, & \text{if } \rho > \hat{\rho} \\ M(-f(K_1^*)\frac{dK_1^*}{d\rho}), & \text{if } \rho < \hat{\rho} \end{cases} \tag{20}$$

It follows from the proof of proposition 1 that $M(-f(K_1^*)\frac{dK_1^*}{d\rho}) < 0$. Therefore, the number of adopters in period 2 must increase as $\rho$ decreases, conditional on the high price regime occurring.

■

# B Propensity-Score Weighting Details

Let $Y^0(i,t)$ represent the emission rate unit $i$ would attain at time $t$ in absence of treatment. Similarly, let $Y^1(i,t)$ represent the emission rate unit $i$ would attain at time $t$ if exposed to the treatment. The effect of the treatment on the outcome for unit $i$ at time $t$ is defined as $Y^1(i,t) - Y^0(i,t)$. Additionally, let $D(i)$ be an indicator function determining if unit $i$ receives the treatment. Also define $P(D=1|X)$ as the propensity score, the probability a unit receives treatment conditional on observed covariates. For this analysis, the treatment group will be all units located in the "challenger" states and the control group will be all other units in CAIR.[38] I use the year 2004 emission rate as the pre-period observation and the 2009 emission rate as the post-period observation. The objective is to estimate the average treatment effect on the treated group (ATT): $E[Y^1(i,2009) - Y^0(i,2004)|D(i)=1]$. Estimation of the ATT requires a weaker assumption on distribution of covariates than would be required to estimate the population average treatment effect (ATE). For identification, require that for all X, $P(D = 1|X) < 1$ , in addition to the unconfoundedness assumption.[39] Appendix D includes marginal kernel density plots of the continuous covariates for each group which demonstrate that this overlap condition is satisfied. Therefore, the average treatment on the treated is given by:

$$
\begin{aligned}
&ATT = E[Y^1(i,2009) - Y^0(i,2009)|D(i)=1] \\
&= \int E[Y^1(i,2009) - Y^0(i,2009)|X(i), D(i)=1] dP(D=1|X) \\
&= \int E[\rho_o \cdot (Y(i,2009) - Y(i,2004))|X(i)] dP(D=1|X) \\
&= E\big[\rho_0 \cdot (Y(i,2009) - Y(i,2004)) \cdot \frac{P(D=1|X)}{P(D=1)}\big] \\
&= E\left[\frac{(Y(i,2009) - Y(i,2004))}{P(D=1)} \cdot \frac{D - P(D=1|X)}{(1-P(D=1|X))}\right]
\end{aligned}
\tag{21}
$$

$$
where \quad \rho_0 = \frac{D - P(D=1|X)}{P(D=1|X)(1-P(D=1|X))}
$$

The third line follows from the unconfoundedness assumption, after controlling for observed covariates, the treatment and control groups would have followed parallel paths absent the intervention. The estimator is the sample analog of the fifth line in (21). Intuitively, the estimator is down weighting the distribution of $Y(i,2009) - Y(i,2004)$ for the untreated group for values of the covariates which are over-represented among the untreated and weighting-up $Y(i,2009) - Y(i,2004)$ for those values of the covariates under-represented among the untreated. I estimate the propensity score using a flexible logit model that includes interactions of all the covariates and quadratic terms.

[38] The semi-parametric DID estimator only allows for one treatment group and one control group so I omit units outside CAIR.

[39] Estimation of the ATE requires the overlap condition: $0 < P(D=1|X) < 1$.

# C Robustness Checks

To address additional identification concerns, I conduct several robustness checks. In Panel A of Table 7 in Appendix D, I estimate the model, excluding 2009 from the sample to account for the possibility that firms reacted quickly to the court decision.[40] Dropping 2009 does not cause any noticeable changes to the estimated effects. In Panel B of Table 7, I restrict the sample to only units within 1000 miles of the centroid of Texas, Florida, or Minnesota. It is possible that the variable "Distance to the Powder River Basin" is not sufficiently controlling for coal purchasing opportunities. For example Florida and New Hampshire may be similar distance to the Powder River Basin but may face significantly different opportunity cost of buying low-sulfur coal. Restricting the sample to only nearby plants does not change the direction or statistical significance of the coefficients of interest.

It is also possible that other political or legal factors are driving differences in pollution abatement and not policy uncertainty. I attempt to address some of these potential concerns in Appendix D Table 8. For instance, during the time frame of this study, some power plants were required to install pollution controls due to the New Source Review (NSR) lawsuits. It is a priori possible that NSR requirements are driving results if many of these lawsuits occurred in CAIR states. Panel A of Table 8 reports estimates of the baseline model on a restricted sample that excludes any plant that was subject to NSR litigation related to $SO_2$ emissions. The results are robust to the exclusion of these plants, which mitigates concerns that NSR lawsuits are impacting the results.

Another potential concern is that other political or institutional factors impacted emission reductions. Panel B in Table 8 shows estimates of the baseline model from equation 4 but only including units in states that had a Republican governor in 2006 and choose PUC chairmen by appointment. In 2006, Texas, Minnesota, and Florida all had Republican governors and appointed PUC chairmen. This restricted sample attempts to deal with possible confounding political factors that would make installing pollution controls more feasible in some states. Since most states did not have both a Republican governor and an appointed PUC commission,[41] 75% of the observation are dropped. However, the point estimate is still positive, statistically significant, and of similar magnitude. This result suggests that the baseline result is not being driven by confounding political factors.

An additional potential concern with the analysis is that the border states' legal challenge initially only challenged the inclusion of Texas plants that were located west of the north-south I-35/I37 corridor. In the baseline analysis, I included all Texas units in the challenger group because EPA almost always levies regulations for entire states to avoid in-state pollution havens. As a robustness check, I have also run the baseline regressions with plants in East

[40] 2009 was after the court ruling so it is possible firms could have reduced emissions after the ruling was made. Scrubbers usually take over a year to install, and coal is usually purchased on one-year contracts so this is unlikely but possible.

[41] Many states elect PUC commissioners.

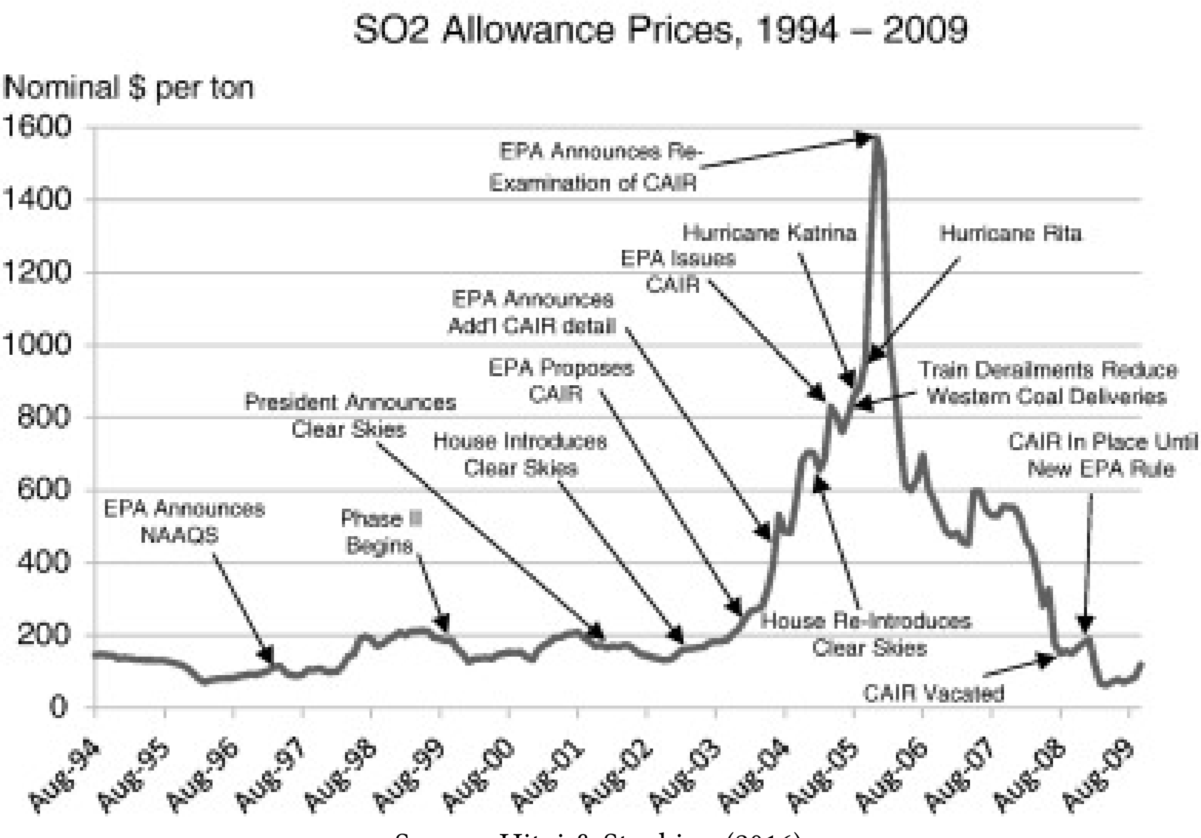
Texas, discluded from the "Challenger" group. The results are shown in Table 10 and the results are robust to this change.

I also account for the possibility that firms were changing electricity output as a method of compliance. In particular, I estimate the model with total annual $SO_2$ emissions in tons as the dependent variable. I also estimate the baseline DID regression with the natural logarithm of $SO_2$ in tons as the dependent variable and also with the emission rate as the outcome variable. The results of all of these regressions are consistent with the baseline model and are presented in Table 12 in the Appendix D.

Finally, I present estimates with alternative standard error clusters. In Table 11, I allow for clustering at the unit level, plant level, state-year level, and state level. These alternative clusters do not change the significance of the estimated effects. In Table 9 also shows that the results are robust to using the entire sample of coal units (including scrubbed units) and also to using an unbalanced panel (including units that did not operate in each year of the sample).

# D Tables and Figures

Figure 7: $SO_2$ Allowance Price History




Source: Hitaj & Stocking (2016)

Table 6: Diff-in-Diff : Acid Rain Program Phase 2 (2000)

| PANEL A: LOG $SO_2$ (LBS/MMBTU) | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Challenger | -0.0541 | -0.0541 | -0.0541 | -0.0541 |
| | (0.0666) | (0.0666) | (0.0671) | (0.0768) |
| CAIR | 0.0272 | 0.0272 | 0.0272 | 0.0272 |
| | (0.0393) | (0.0393) | (0.0396) | (0.0453) |
| Controls | Yes | Yes | Yes | No |
| Year_FE | No | Yes | Yes | Yes |
| State_FE | No | No | Yes | No |
| Unit_FE | No | No | No | Yes |
| N | 2704 | 2704 | 2704 | 2704 |
| r2 | 0.874 | 0.874 | 0.881 | 0.921 |

| PANEL B: DECISION TO INSTALL SCRUBBER | (1) | (2) |
|---|---|---|
| Challenger | 0.00463 | 0.342 |
| | (0.0266) | (0.359) |
| CAIR | -0.0548** | -0.602** |
| | (0.0218) | (0.301) |
| Model | OLS | Probit |
| Controls | Yes | Yes |
| N_Challenger | 46 | 46 |
| N | 676 | 676 |
| r2 | 0.0714 | |

Panel A reports regressions results for a Diff-in-Diff regression with 1996-1999 as the pre-period and 2000-2001 as the post-period and $SO_2$ emission rate as the dependent variable. For Panel A, standard errors are clustered at the unit level. Panel B reports estimates for both OLS and Probit models where the dependent variable is a binary decision to install a Scrubber by 2001, the sample includes all units that did not already have a scrubber installed in 1996, * indicates $p < 0.10$, ** indicates $p < 0.05$, and *** indicates $p < 0.01$

Table 7: Robustness Checks - Diff-in-Diff with Sample Restrictions 1 - Dep Var: Log SO2 (lbs. per MMBtu)

PANEL A: DROP 2009

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Challenger | 0.0930*** | 0.0930*** | 0.0930*** | 0.0930*** |
| | (0.0298) | (0.0298) | (0.0299) | (0.0321) |
| CAIR | -0.131*** | -0.131*** | -0.131*** | -0.131*** |
| | (0.0257) | (0.0257) | (0.0258) | (0.0277) |
| Controls | Yes | Yes | Yes | No |
| Year_FE | No | Yes | No | Yes |
| State_FE | No | No | Yes | No |
| Unit_FE | No | No | No | Yes |
| N | 4669 | 4669 | 4669 | 4669 |
| r2 | 0.702 | 0.716 | 0.712 | 0.797 |

PANEL B: ONLY PLANTS NEAR MN,FL,TX

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Challenger | 0.113*** | 0.113*** | 0.113*** | 0.113*** |
| | (0.0404) | (0.0405) | (0.0405) | (0.0432) |
| CAIR | -0.197*** | -0.198*** | -0.197*** | -0.198*** |
| | (0.0331) | (0.0332) | (0.0332) | (0.0354) |
| Controls | Yes | Yes | Yes | No |
| Year_FE | No | Yes | No | Yes |
| State_FE | No | No | Yes | No |
| Unit_FE | No | No | No | Yes |
| N | 3327 | 3327 | 3327 | 3327 |
| r2 | 0.510 | 0.538 | 0.528 | 0.668 |

Panel A reports regressions results for the baseline Diff-in-Diff regression from equation 4 excluding data from 2009. For this regression, the pre-period is 2002-2004 and 2005-2008 is the post-period. The court made a ruling in December 2008, so it is possible that firms could react to the announcement by reducing emissions in 2009. Panel B estimates the baseline model from equation 4 but excluding all units that are located further than 1000 miles from the centroid of TX, MN, or FL. All standard errors are clustered at the unit level, * indicates $p < 0.10$, ** indicates $p < 0.05$, and *** indicates $p < 0.01$

Table 8: Robustness Checks - Diff-in-Diff with Sample Restrictions 2 - Dep Var: Log SO2 (lbs. per MMBtu)

| Panel A: Drop Units Subject to NSR Lawsuits | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Challenger | 0.107*** | 0.107*** | 0.107*** | 0.107*** |
| | (0.0369) | (0.0370) | (0.0371) | (0.0395) |
| CAIR | -0.138*** | -0.138*** | -0.138*** | -0.138*** |
| | (0.0308) | (0.0309) | (0.0310) | (0.0330) |
| Controls | Yes | Yes | Yes | No |
| Year_FE | No | Yes | No | Yes |
| State_FE | No | No | Yes | No |
| Unit_FE | No | No | No | Yes |
| N | 5007 | 5007 | 5007 | 5007 |
| r2 | 0.609 | 0.633 | 0.629 | 0.741 |

| Panel B: Only Units in States with Republican Appointed PUC Chairmen | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Challenger | 0.155** | 0.155** | 0.155* | 0.155* |
| | (0.0785) | (0.0787) | (0.0787) | (0.0839) |
| CAIR | -0.223*** | -0.223*** | -0.223*** | -0.223** |
| | (0.0829) | (0.0831) | (0.0831) | (0.0886) |
| Controls | Yes | Yes | Yes | No |
| Year_FE | No | Yes | No | Yes |
| State_FE | No | No | Yes | No |
| Unit_FE | No | No | No | Yes |
| N | 1376 | 1376 | 1376 | 1376 |
| r2 | 0.578 | 0.601 | 0.588 | 0.709 |

Panel A reports regressions results for the baseline Diff-in-Diff regression from equation 4 excluding any unit that were subject to New Source Review litigation related to $SO_2$ emissions. During the time frame of this study, some power plants were required to install pollution controls due to NSR regulations, all of these plants are excluded from the sample. I thank Ian Lange for supplying NSR information. Panel B estimates the baseline model from equation 4 but only including units in states that had a Republican governor in 2006 and choose PUC chairmen by appointment. In 2006, TX, MN, and FL all had Republican governors and appointed PUC chairmen. This restricted sample attempts to deal with possible confounding political factors that would make installing pollution controls more feasible in some states. All standard errors are clustered at the unit level. * indicates $p < 0.10$, ** indicates $p < 0.05$, and *** indicates $p < 0.01$

Table 9: Robustness Checks 3 - Diff-in-Diff with Sample Changes - Dep Var: Log SO2 (lbs. per MMBtu)

Panel A: Unbalanced Panel

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Challenger | 0.142*** | 0.138*** | 0.141*** | 0.127*** |
| | (0.0321) | (0.0317) | (0.0326) | (0.0334) |
| CAIR | -0.130*** | -0.131*** | -0.122*** | -0.113** |
| | (0.0361) | (0.0364) | (0.0409) | (0.0459) |
| Controls | Yes | Yes | Yes | No |
| Year_FE | No | Yes | No | Yes |
| State_FE | No | No | Yes | No |
| Unit_FE | No | No | No | Yes |
| N | 6329 | 6329 | 6329 | 6464 |
| r2 | 0.603 | 0.623 | 0.618 | 0.757 |

Panel B: All Units (Including Scrubbed)

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Challenger | 0.104*** | 0.103*** | 0.103*** | 0.0959*** |
| | (0.0319) | (0.0317) | (0.0319) | (0.0338) |
| CAIR | -0.0558* | -0.0568* | -0.0539 | -0.0457 |
| | (0.0326) | (0.0328) | (0.0333) | (0.0379) |
| Controls | Yes | Yes | Yes | No |
| Year_FE | No | Yes | No | Yes |
| State_FE | No | No | Yes | No |
| Unit_FE | No | No | No | Yes |
| N | 8096 | 8096 | 8096 | 8278 |
| r2 | 0.789 | 0.798 | 0.795 | 0.865 |

Panel A reports regressions results for the baseline Diff-in-Diff regression from equation 4 but includes all units even units that were not operating in each time period throughout the sample. Panel B reports regressions results for the Diff-in-Diff regression from equation 4 but includes the full population of coal units including units that already had scrubbers installed before 2005. $^{*}$ indicates $p < 0.10$, $^{**}$ indicates $p < 0.05$, and $^{***}$ indicates $p < 0.01$

Table 10: Robustness Check - Disclude East Texas Plants from Challenger Group

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Challenger | 0.134*** | 0.134*** | 0.134*** | 0.134*** |
| | (0.0371) | (0.0371) | (0.0372) | (0.0397) |
| CAIR | -0.165*** | -0.165*** | -0.165*** | -0.165*** |
| | (0.0311) | (0.0311) | (0.0312) | (0.0332) |
| Controls | Yes | Yes | Yes | No |
| Year_FE | No | Yes | No | Yes |
| State_FE | No | No | Yes | No |
| Unit_FE | No | No | No | Yes |
| N | 5335 | 5335 | 5335 | 5335 |
| r2 | 0.579 | 0.606 | 0.596 | 0.721 |

This table reports regressions results for the Diff-in-Diff regression from equation 4 but but does not include plants in East Texas in the Challenger Group. * indicates $p < 0.10$, ** indicates $p < 0.05$, and *** indicates $p < 0.01$

Table 11: Alternative Std. Error Clusters: DiD Dep Var: Log SO2 (lbs. per MMBtu)

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Challenger | 0.134*** | 0.134*** | 0.134*** | 0.134*** |
| | (0.0397) | (0.0514) | (0.0431) | (0.0455) |
| CAIR | -0.165*** | -0.165*** | -0.165*** | -0.165*** |
| | (0.0332) | (0.0467) | (0.0364) | (0.0519) |
| Year_FE | Yes | Yes | Yes | Yes |
| Unit_FE | Yes | Yes | Yes | Yes |
| SE_Cluster | Unit | Plant | State-Year | State |
| N | 5335 | 5335 | 5335 | 5335 |
| r2 | 0.721 | 0.721 | 0.721 | 0.721 |

This table reports regressions results for the baseline Diff-in-Diff regression from equation 4 allowing for alternative standard error clustering by Unit, Plant, State-Year, and State. * indicates $p < 0.10$, ** indicates $p < 0.05$, and *** indicates $p < 0.01$

Table 12: Robustness Checks - Alternative Dependent Variables

Panel A: Levels $SO_2$ (tons)

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Challenger | 755.9* | 759.4* | 771.6* | 677.4 |
| | (437.9) | (437.4) | (433.8) | (500.2) |
| CAIR | -1808.5*** | -1810.4*** | -1812.5*** | -1787.9*** |
| | (246.4) | (246.5) | (246.1) | (271.7) |
| Controls | Yes | Yes | Yes | No |
| Year_FE | No | Yes | No | Yes |
| State_FE | No | No | Yes | No |
| Unit_FE | No | No | No | Yes |
| N | 5335 | 5335 | 5335 | 5335 |
| r2 | 0.798 | 0.812 | 0.804 | 0.869 |

Panel B: Log($SO_2$ (tons))

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Challenger | 0.134*** | 0.134*** | 0.134*** | 0.131** |
| | (0.0371) | (0.0371) | (0.0372) | (0.0620) |
| CAIR | -0.169*** | -0.169*** | -0.169*** | -0.239*** |
| | (0.0311) | (0.0311) | (0.0312) | (0.0357) |
| Controls | Yes | Yes | Yes | No |
| Year_FE | No | Yes | No | Yes |
| State_FE | No | No | Yes | No |
| Unit_FE | No | No | No | Yes |
| N | 5335 | 5335 | 5335 | 5335 |
| r2 | 0.826 | 0.837 | 0.833 | 0.829 |

Panel C: $SO_2$ (lbs/MMBtu)

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Challenger | 0.133*** | 0.133*** | 0.133*** | 0.133*** |
| | (0.0309) | (0.0309) | (0.0310) | (0.0330) |
| CAIR | -0.165*** | -0.166*** | -0.165*** | -0.166*** |
| | (0.0258) | (0.0258) | (0.0258) | (0.0275) |
| Controls | Yes | Yes | Yes | No |
| Year_FE | No | Yes | No | Yes |
| State_FE | No | No | Yes | No |
| Unit_FE | No | No | No | Yes |
| N | 5335 | 5335 | 5335 | 5335 |
| r2 | 0.799 | 0.808 | 0.808 | 0.874 |

This table reports regressions results for the baseline Diff-in-Diff regression from equation 4 with alternative dependent variables. Panel A uses the level of emissions, $SO_2$ in tons, Panel B uses the natural logarithm of $SO_2$ in tons as the outcome variable, and Panel C uses the emission rate of $SO_2$ in lbs/MMBtu. All standard errors are clustered at the unit level. * indicates $p < 0.10$, ** indicates $p < 0.05$, and *** indicates $p < 0.01$

Figure 8: Marginal Kernel Density Plots for Observed Covariates