Online Appendix

CAN POLLUTION MARKETS WORK IN DEVELOPING COUNTRIES? EXPERIMENTAL EVIDENCE FROM INDIA

MICHAEL GREENSTONE

ROHINI PANDE

NICHOLAS RYAN

ANANT SUDARSHAN

This online appendix contains five different parts. Appendix A gives more information on the experimental design. Appendix B describes the data sources and cleaning. Appendix C covers the Continuous Emissions Monitoring System (CEMS) data on pollution, specifically, including imputation rules for missing pollution data. Appendix F gives additional empirical results to support those in the main text. Appendix G provides our benefit-cost analysis.

A. APPENDIX: EXPERIMENTAL DESIGN

This section gives more information about the experimental design. Table A1 gives the timeline of the experimental intervention and data collection. Table A2 describes the duration and market cap for each compliance period of the emissions market. Table A3 describes attrition in the sample by treatment arm and with respect to each source of data. Table A4 duplicates the balance table in the main text but without the sample restriction to plants that report CEMS data.

Table A1. Intervention timeline

	Compliance Period	Data Collection	
		Survey	CEMS
Dec-2018		Baseline survey	
Apr-2019			CEMS data begins
Jul-2019	Mock-I		
Aug-2019	Mock-II		
Sep-2019	Period-I		
Oct-2019	Period-II		
Nov-2019	Period-III		
Jan-2020	Period-IV		
Feb-2020	Period-V		
Mar-2020	Period-VI		
Apr-2020	Interregnum *(COVID-19)*		
Oct-2020	Mock-III		
Nov-2020	Interregnum *(Diwali)*	Endline survey	
Dec-2020	Period-VII		
Jan-2021	Period-VIII		
Feb-2021	Period-IX		
Mar-2021	Period-X		

Compliance periods were of heterogeneous length, though most lasted approximately one month; of particular note, Period-III began in the middle of November and lasted 45 days until early January. Baseline and endline surveys collected data on plant and boiler house costs, revenue, and emissions abatement mechanisms. While CEMS device readings were collected from April 2019 onward, data availability was low until the emissions trading scheme commenced in July 2019. During mock periods, plants simulated live period transactions with monetary vouchers. We had two interregnum periods where the market was closed: the first wave of the COVID-19 pandemic and shutdowns, and Diwali in 2020. Plant production remained sufficiently high during Diwali in 2019 to continue market operations.

Table A2. Compliance periods and market caps

Period	Start Date	End Date	Days	Cap (kg/30 days)	Per-plant Cap (kg/30 days)	Total Cap (kg)
Mock-I	2019/07/15	2019/08/12	29	280,000	1,728	270,667
Mock-II	2019/08/13	2019/09/15	34	280,000	1,728	317,333
Compliance-I	2019/09/16	2019/10/15	30	280,000	1,728	280,000
Compliance-II	2019/10/16	2019/11/15	31	200,000	1,235	206,667
Compliance-III	2019/11/16	2019/12/31	46	180,000	1,111	276,000
Compliance-IV	2020/01/01	2020/01/31	31	170,000	1,049	175,667
Compliance-V	2020/02/01	2020/02/29	29	170,000	1,049	164,333
Compliance-VI	2020/03/01	2020/03/21	21	170,000	1,049	119,000
Interregnum-I	2020/03/22	2020/10/11	204	-	-	-
Mock-III	2020/10/12	2020/11/11	31	170,000	1,049	175,667
Interregnum-II	2020/11/12	2020/11/30	19	-	-	-
Compliance-VII	2020/12/01	2020/12/31	31	170,000	1,049	175,667
Compliance-VIII	2021/01/01	2021/01/31	31	170,000	1,049	175,667
Compliance-IX	2021/02/01	2021/02/28	28	170,000	1,049	158,667
Compliance-X	2021/03/01	2021/03/31	31	170,000	1,049	175,667

This table reports the start and end date of compliance periods and the market cap of each period. The market cap is the total amount of PM emissions – summed up across all market participants - that is allowed *per month (30 days)* under the Emissions Trading scheme. The total market cap varies across compliance periods, due to the duration of the compliance period. Specifically, the total market cap in a compliance period is the market cap $\times$ 30 / (number of days in the compliance period). The per-plant cap is calculated by dividing the market cap by 162, the number of in-sample plants in the treatment arm. The market was closed during Interregnum-I due to the COVID-19 pandemic and during Interregnum-II following the Divali festival.

Table A3. Sample determination and attrition by treatment status

	Control	Treatment	Total
Plants that received treatment assignment	168	174	342
Closed/extinct plants with treatment assignment	10	10	20
Operational-at-baseline plants with treatment assignment	158	164	322
Plants removed from ETS sample by GPCB	2	2	4
In-sample plants	**156**	**162**	**318**
Plants incompetely treated due to closure	7	6	13
Plants completely treated	149	156	305
In-sample plants surveyed at ETS Baseline	147	157	304
In-sample plants manually stack sampled at ETS Baseline	147	157	304
In-sample plants with GPCB administrative data	156	162	318
In-sample plants reporting CEMS data	**136**	**156**	**292**
In-sample plants surveyed at ETS Endline	142	153	295
Treated plants with market trading data	-	155	155

This table reports the sample determination and attrition during the ETS experiment. Of the original ETS-CEMS sample of 373 plants, 342 operational plants received treatment assignment in May 2019 (row 1). Of these 342 plants included in the ETS treatment randomization, 20 plants were extinct or permanently closed (row 2). The permanent shutdown status of these 20 plants has been verified with Ringelmann survey panel data covering the sample from March 2018 to June 2019, as well as regulatory inspection and audit documentation on the GPCB administrative portal. The 342 plants that received treatment assignment, less the 20 plants that received assignment while extinct or shutdown, yield 322 operational plants with treatment assignment at baseline (row 3). Four of these 322 operational-at-baseline plants were officially removed from the ETS sample by GPCB after the treatment assignment (row 4). Three of the removed plants (2 in control, 1 in treatment) are seasonal sugar cooperatives, operational for only four months of the year; the fourth treatment plant is a particle-board producing plant which uses bagasse, rather than coal, as fuel. Of the 318 in-sample plants, 13 are known to have been incompletely treated by the intervention, due to temporary financial closure before or after the treatment assignment was done (row 6). The 304 plants surveyed at baseline are distinct from the 304 plants manually sampled, and are therefore reported separately (rows 8, 9). This paper reports experimental results from the sample of 292 plants reported at least one day of CEMS data from April 16, 2019 to April 3rd, 2021 (row 11). Of the 162 in-sample plants in the treatment group, 153 plants have market trading data (row 13).

Table A4. Balance of plant characteristics by treatment status, full sample

	Treatment (1)	Control (2)	Difference (3)
Panel A: Plant Measures			
Total electricity cost (1,000 USD)	456.2	389.1	67.1
	[853.1]	[660.7]	(89.6)
Log(plant total heat output)	15.6	15.5	0.085
	[0.61]	[0.59]	(0.067)

Size as recorded on environment consent (1 to 3)	1.36	1.40	-0.038
	[0.63]	[0.65]	(0.073)
Small-scale (size=1)	0.72	0.69	0.033
	[0.45]	[0.47]	(0.053)
Large-scale (size=3)	0.083	0.088	-0.0056
	[0.28]	[0.28]	(0.032)
Number of stacks	1.08	1.05	0.035
	[0.41]	[0.21]	(0.037)
Textiles sector (=1)	0.85	0.85	-0.0032
	[0.36]	[0.36]	(0.041)
Gross Sales Revenue in 2017 (1,000 USD)	12614.5	13628.3	-1013.8
	[42698.0]	[53258.9]	(5680.6)
Panel B: Plant Abatement and Investment Cost			
Boiler house employment	36.8	31.7	5.13
	[32.5]	[30.0]	(3.59)
Boiler house capital expenditure (1,000 USD)	198.3	164.2	34.0
	[398.6]	[190.9]	(36.7)
Boiler house operating cost (1,000 USD)	138.1	111.0	27.1
	[202.6]	[84.9]	(17.6)
APCD: Cyclone present	0.98	0.97	0.0081
	[0.14]	[0.16]	(0.017)
APCD: Bag filter present	0.80	0.86	-0.055
	[0.40]	[0.35]	(0.043)
APCD: Scrubber present	0.64	0.61	0.032
	[0.48]	[0.49]	(0.056)
APCD: ESP present	0.11	0.082	0.033
	[0.32]	[0.27]	(0.034)
Panel C: Plant Pollution Measures			
Plant total PM mass rate (kg/hr)	3.62	3.51	0.11
	[4.86]	[3.76]	(0.50)
Plant mean PM concentration (mg/Nm^3)	177.9	168.5	9.37
	[153.6]	[151.5]	(17.5)
Plant mean Ringelmann score (1 to 5)	1.36	1.35	0.0090
	[0.42]	[0.37]	(0.045)
Above regulatory standard at ETS baseline (=1)	0.33	0.28	0.052
	[0.47]	[0.45]	(0.053)
Number of plants	162	156	

This table shows differences in plant measures (panel A), plant abatement and investment cost (panel B), and plant pollution (panel C) between the treatment and control groups of plants in the baseline survey conducted from December 2018 to January 2019. This sample consists of 318 plants in the ETS experiment. In panel B, cyclone, bag filter, scrubber, and electrostatic precipitator (ESP) are different devices used to reduce emissions. Some plants did not respond to some questions in the survey. For the control group, the numbers of observations are 137 for boiler house capital expenditure, 141 for gross sales revenue, 148 for the Ringelmann score, 156 for plant total heat output, and 147 for the rest. For the treatment group, the numbers of observations are 147 for boiler house capital expenditure, 150 for gross sales revenue, 160 for Ringelmann score, 162 for plant total heat output and the number of stacks, and 157 for the rest. The first and second columns show means with standard deviations given in brackets. The third column shows the coefficient from regressions of each variable on treatment, with robust standard errors in parentheses. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

B. APPENDIX: DATA

This Appendix B discusses data from our plant survey and from administrative records on permit trade. The following Appendix C covers data from Continuous Emissions Monitoring Systems (CEMS).

B.1. Survey data

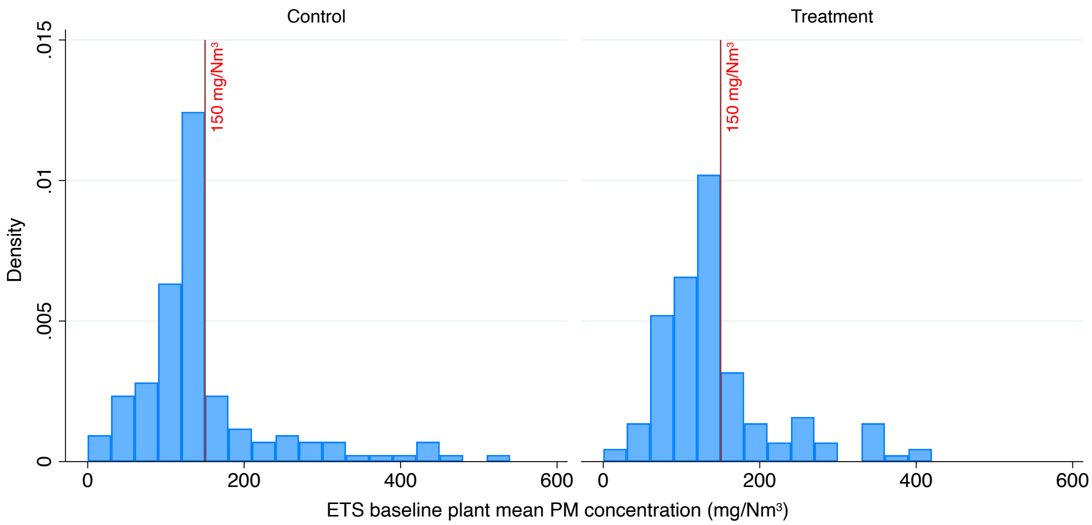
The ETS baseline survey was conducted from December 2018 to February 2019. The unit of analysis is a plant. The survey consists of three main sections: a general section, a technical section, and a pollution sampling section. In the general section, researchers at J-PAL South Asia asked the plant managers questions about plant operations. Researchers then spoke to boiler engineers to collect information about the machinery specifications for the technical section.

As part of the technical survey environmental labs collected samples from the stack(s) (i.e., chimney) attached to the boiler and/or thermopack to measure the PM concentration and PM mass rate. Of 304 plants covered in the technical survey 289 have only a single stack. Participation in the survey is voluntary. Plants were notified by J-PAL South Asia that their name and data would not be published in any report, and their data would never be shown to the Gujurat Pollution Control Board (GPCB). J-PAL covered the cost of stack sampling and surveys.

Figure B1 shows the distribution of emissions concentrations in the baseline survey by treatment arm. The red vertical lines at 150 mg/Nm^3 indicate the regulatory standard. Many plants are out of compliance with the standard, sometimes by a factor of two or more. Table I, panel C shows the plant's mean PM concentration from sampling and an indicator for non-compliance. A total of 66% of plants are in compliance at baseline in the treatment group and 72% in the control group.

In addition to stack sampling, J-PAL South Asia had conducted ten rounds of Ringelmann surveys from February 2018 to June 2019. The Ringelmann score is a scale for measuring the density of smoke as it appears to the naked eye. The scale has five levels of density. Score 1 to 5 correspond to an opacity of 20%, 40%, 60%, 80% and 100%. Prior to Ringelmann surveys, GPCB informed plants that the information collected would not be used for determining compliance with

Figure B1. Distribution of Pollution Before the Experiment



This figure shows the distributions of the plant PM concentration by treatment status as measured by manual iso-kinetic stack sampling at the ETS baseline (December 2018 to January 2019). One PM sample was collected from each industrial stack by a third-party laboratory. The histograms are truncated at the 95th percentile (520 mg/Nm^3). The red, vertical lines indicate the regulatory concentration standard of 150 mg/Nm^3. At the ETS baseline, 28% of sampled plants in the control group and 34% of sampled plants in the treatment group had readings above this standard.

the GPCB norms or any other regulatory purpose.

In Table I variables in panel A are from the general section of the ETS baseline survey, and those in Panel B are from the technical section. In panel B, cyclones, bag filters, scrubbers, and electrostatic precipitators (ESPs) are air pollution control devices (APCDs) used to abate PM emissions. In panel C, the plant's total PM mass rate is the sum of the plant's stacks' PM mass rates measured from stack sampling. The plant's mean Ringelmann score is the average of scores from the four pre-treatment rounds of Ringelmann surveys conducted from April 2019 to June 2019.

B.2. Trading data

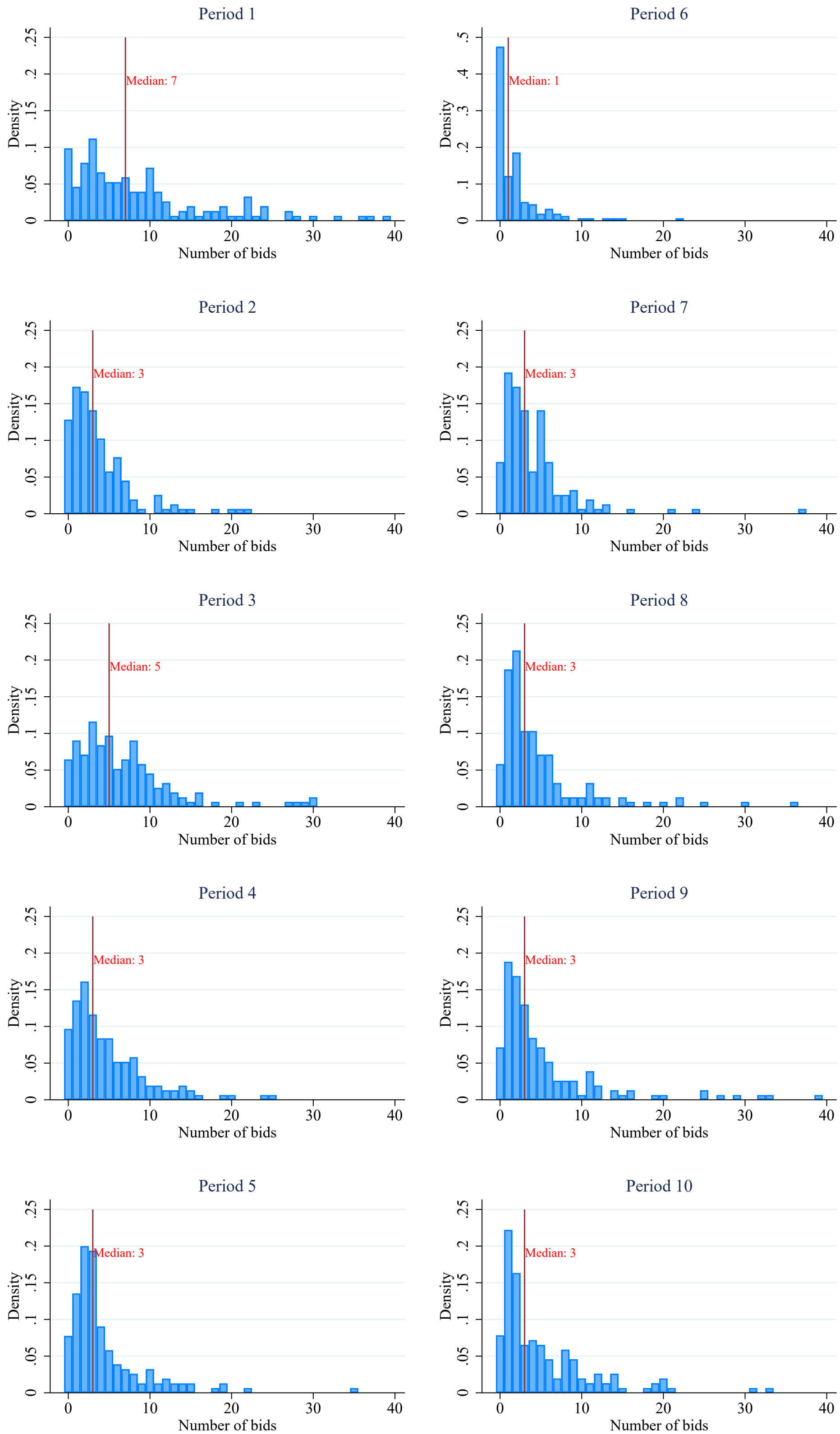
The paper uses administrative data on permit bids and offers from NeML, the market operator. Table B1 shows summary statistics on permit bids (panel A) and executed trades (panel B). Figure B2 shows the distribution of the number of bids placed per plant in each compliance period.

Table B1. Trading data summary statistics

	All	Purchase	Sale
Panel A: Order			
Order quantity (kg)	411.61	429.50	398.78
	(707.98)	(565.09)	(794.52)
Order price (Rs/kg)	11.25	9.47	12.52
	(11.56)	(10.50)	(12.10)
Order price (Rs/kg), weighted by quantity	9.23	8.42	9.86
	(8.49)	(8.71)	(8.27)
Observations	8433	3520	4913
Panel B: Trade			
Trade quantity (kg)	360.42	389.36	327.18
	(564.45)	(544.09)	(585.37)
Trade price (Rs/kg)	9.29	9.19	9.40
	(7.39)	(9.31)	(4.20)
Trade price (Rs/kg), weighted by quantity	8.40	8.16	8.73
	(6.17)	(7.26)	(4.21)
Observations	3775	2018	1757

This table shows the mean of order quantity and price (panel A) and trade quantity and price (panel B), with the standard deviation given in the brackets.

Figure B2. Distribution of number of bids placed per plant by compliance period



This figure presents the distributions of number of bids placed per plant by compliance period, truncated at 40 (about 99th percentile). The bin width is 1. The red line indicates the median number of bids placed.

C. APPENDIX: POLLUTION MONITORING

C.1. Measuring plant emissions

We describe how we construct plant-level monthly average PM mass (in kilograms). CEMS provides stack-level daily reporting hours and uncalibrated daily average PM mass rate (kg/hr) or PM concentration (mg/Nm^3). A plant might have multiple stacks. A month in our analysis is defined as the 16th of this month to the 15th of next month.[32] We follow four steps: calibration, truncation, imputation and aggregation.

1 *Calibration.* The raw data set consists of 242,303 daily observations of 337 stacks (318 plants) from April 16th, 2019 to April 3rd, 2021. Plants can choose what kind of CEMS device to install (Type 1 or Type 2), so long as the device meets technical standards and passes the calibration test. Generally the devices calibrated to mass rates (Type 1) are simpler and less expensive than having a combination of a concentration device and a flowmeter to measure volume. Most plants therefore install Type 1 devices, without any mandate to do so from the regulator. The Type-1 devices measure the daily average PM mass rates (kg/hr), and the Type-2 devices measure the daily average PM concentration (mg/Nm^3).

The PM mass rate and concentration are calibrated according to the device type. For a stack i (j) that uses Type-1 (2) devices, we calibrate its average PM mass rate (concentration) on the day d using the formula

$$\text{PM\_Rate}_{i,d} = m_i \text{PM\_Rate}^{raw}_{i,d} + c_i,$$

$$\text{PM\_Conc}_{j,d} = m_j \text{PM\_Conc}^{raw}_{j,d} + c_j,$$

where m and c are stack's calibration factors. Any negative calibrated value is set to missing. We

[32]This definition was an artifact of the market's initial timing. The first compliance period began on the 16th of September, following two mock periods running from the 16th of July. GPCB intended to start on the 1st of July but pushed the market back by two weeks to allow a grace period and increase CEMS reporting at the start of the market.

convert the mass rate to concentration, or vice versa, using

$$\text{PM\_Conc}_{i,d}^{cal} = \frac{1000^2 \ \text{PM\_Rate}_{i,d}^{cal}}{(3600 \ \text{max\_velocity}_i) \times \text{stack\_area}_i},$$

where max_velocity is the maximum flue velocity (m/s) of calibration samples, and stack_area is the stack cross-sectional area (m^2).

C.2. Imputation

Before we do any imputation, we truncate outliers in the non-missing data. Within each stack-day we take the average of the reported PM kg/hr at the minute-level. This gives us a set of stack-days across the entire sample. We then set to missing all those stack-days which fall above the 99th percentile.

Our final unit of observation is stack-month emissions (which are then summed within plants to get plant-months). To calculate emissions at the stack-month level we sum the emissions at the minute-level across all the minutes in the stack-month. Thus, our imputation aims to fill in all missing stack-minutes. We impute in two stages: first within the stack-week observation unit, and second within or across stack-months.

The first step is within stack-week imputation. Emissions are imputed for all missing stack-minutes in stack-weeks which have any data reported. To do this, we fill in all minutes in each week with the stack's average reported emissions during that week.

The second step is to either perform within stack-month imputation or across stack-month imputation. How we do this depends on which imputation rule we use. The three options are as follows:

- **The "no imputation" rule** only performs within stack-month imputation. Any remaining missing minutes from the first step (i.e. those in months with some data reported but in weeks with no data reported) are imputed at the average reported value across the entire month. This procedure does not perform any across stack-month imputation and leaves stack-months with no reported data as missing.

- **Rule A ("Stack Experiment")** imputes stack-weeks which do not report any data at the average level of the reported values for that stack over the entire ETS experiment (excluding the interregnum periods).

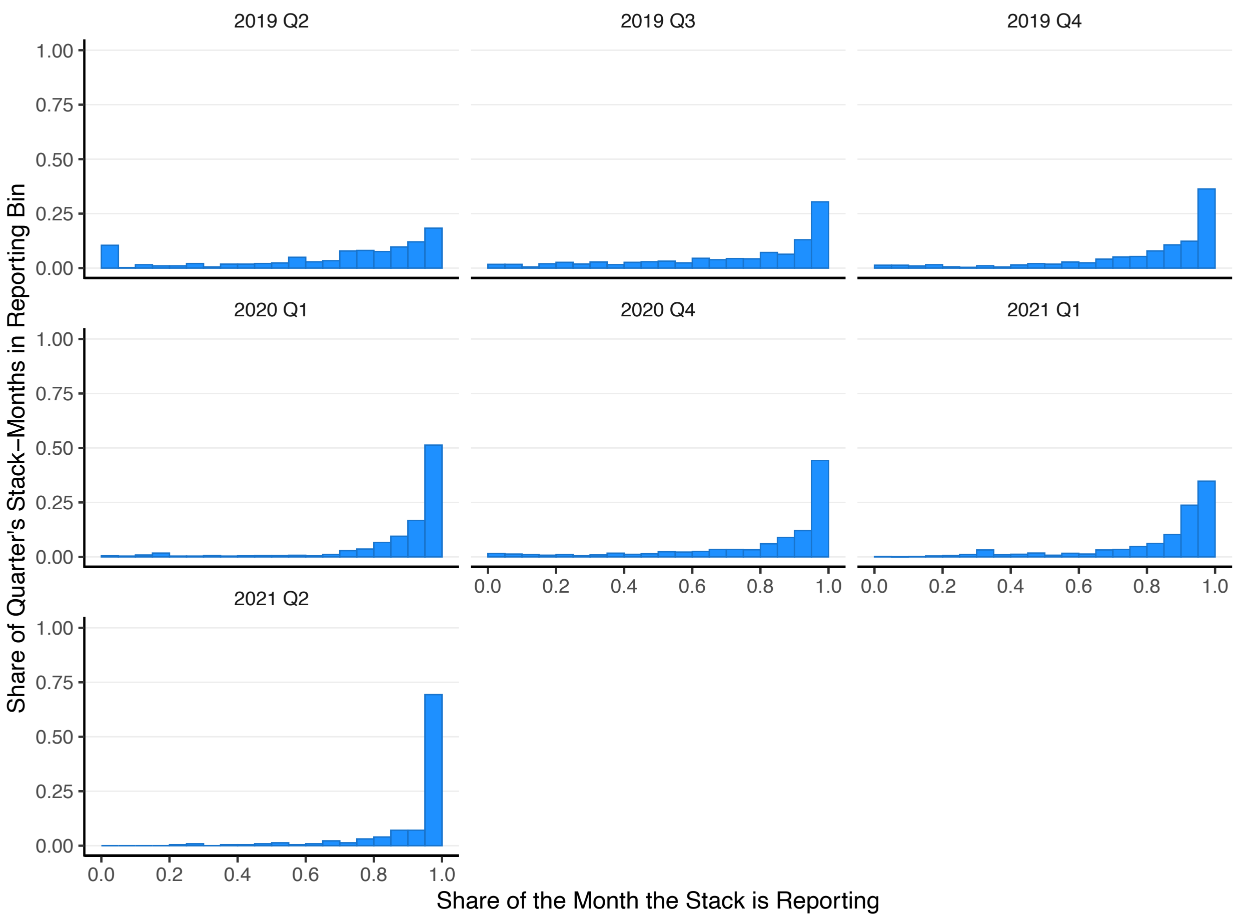
- **Rule B ("Treatment")** imputes stack-weeks which do not report any data at the average level of the reported values for that appropriate Treatment/Control group for that month (i.e. if that stack is part of the control group we average over just control plants in that month).

In Figure C1 we show the distribution of reporting frequency among those stack-months which require some intra-month imputation.

C.3. Treatment effect on pollution under alternate imputation rules

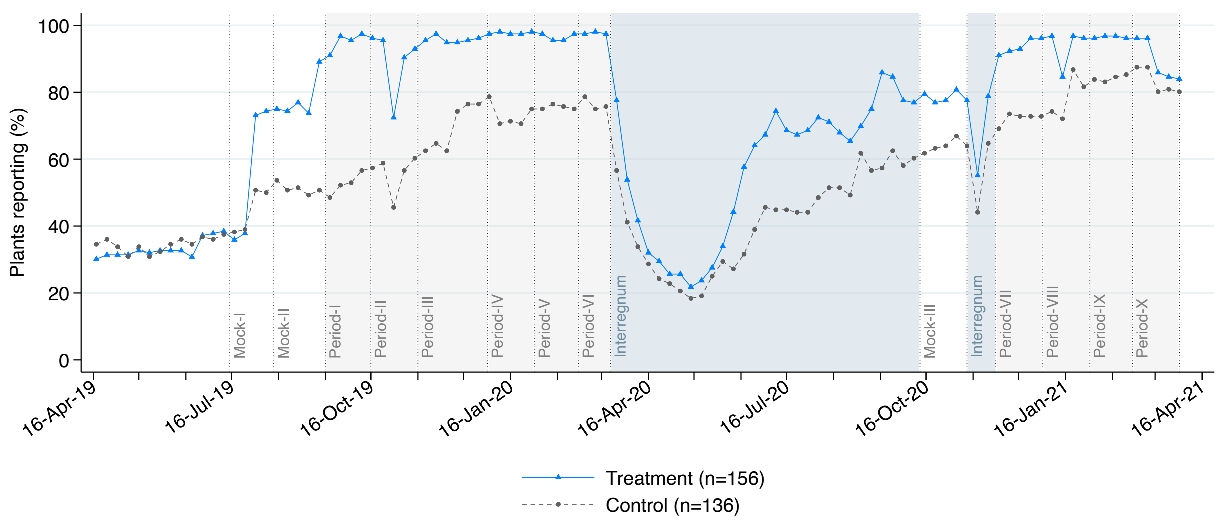
Figure V shows the time series of pollution by treatment status. Here we show the same series under alternate imputation rules for missing data. Figure C2 shows the data availability of CEMS data on pollution by treatment arm. Figure C3 replicates Figure V, from the main text, but with alternate imputation rules for missing data. Table C1 summarizes the level of log PM emissions by imputation rule and Figure C4 compares the distribution of pollution under different rules.

Figure C1. Share of data available within month for months with partial data



This figure plots a histogram of data availability at the stack-month level. The only stack-months included are those which require intra-month imputation, which are those with some, but not complete, minute-level reporting throughout the month. This represents 73% of the stack-months in the sample. The y-axis represents the portion of all plant-months in that panel which fall into the corresponding bin. Each panel represents a different quarter-year of the sample, excluding interregnum periods.

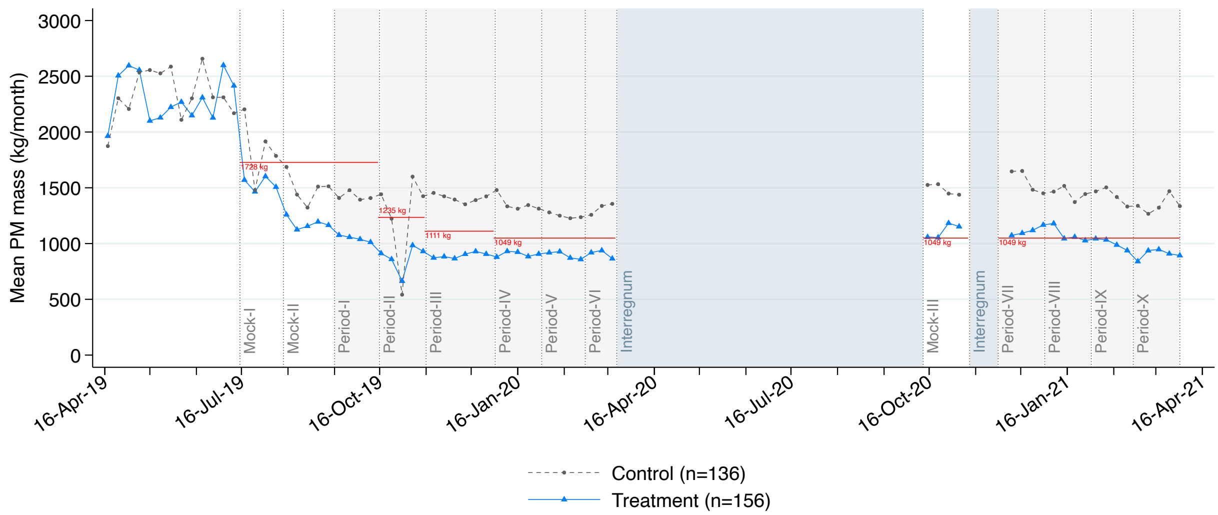
Figure C2. Data availability from CEMS by treatment status



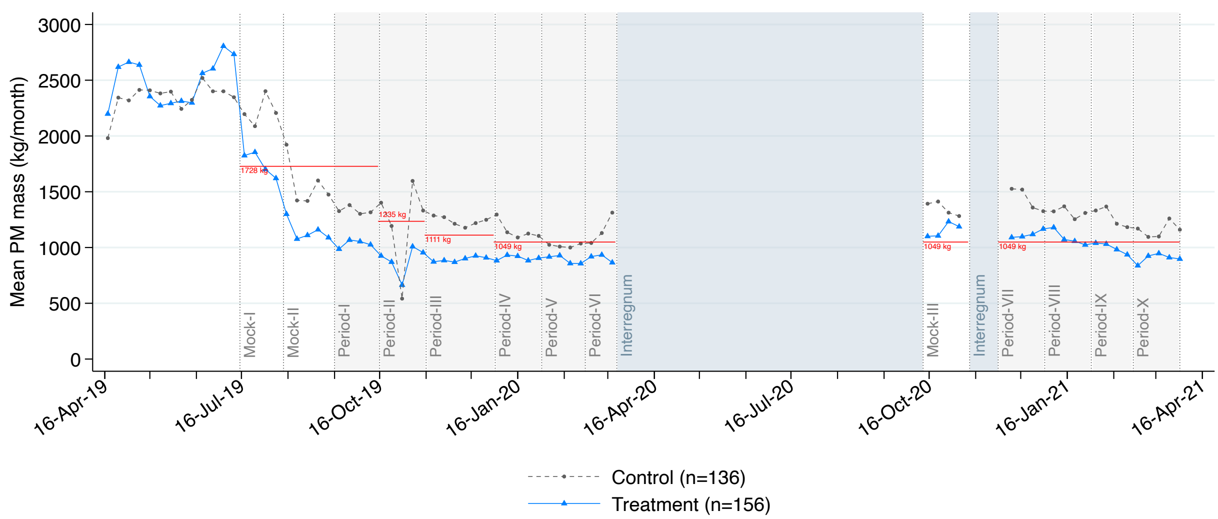
The figure shows the percentage of plants reporting, at weekly frequency, from April 2019 to March 2021. The missing pollution readings are imputed within a stack-week, but not across stacks or weeks. This sample consists of 292 plants that had at least one day of PM data from CEMS devices during the ETS experiment. The treatment group is represented by the solid (blue) line, and control group by the dashed (grey) line. The grey regions mark the ten compliance periods in the emissions market. The light blue regions mark the two interregnum periods when the emissions market was closed.

Figure C3. PM emissions by treatment status

A. Rule A: Stack-Experiment



B. Rule B: Treatment-Month



The figure shows the weekly mean per-plant PM emissions in kilograms calculated at a monthly rate equivalent, from April 2019 to March 2021. In the top panel, the missing pollution readings are imputed within stack-week, and then within stack-experiment. In the bottom panel, they are imputed within stack-week, and then within treatment-month. Appendix C.1 provides a detailed note on the construction of the PM emission variable. This sample consists of 292 plants that had at least one day of PM data from CEMS devices during the ETS experiment. The treatment group is represented by the solid (blue) line, control group by the dashed (grey) line. The grey regions mark the ten compliance periods in the emissions market. The light blue regions mark interregnum periods when the emissions market was closed. The horizontal (red) lines denote the per-plant month market cap for each period. The aggregate market caps for each compliance period were: 280 tons per 30 days (for Mock-I, Mock-II, and Period-I), 200 tons per 30 days (for Period-II), 180 tons per 30 days (for Period-III), and 170 tons per 30 days thereafter.

Table C1. Mean of the log(PM emissions) by imputation rules

	Control	Treatment	All
No Imputation	6.67	6.52	6.58
	[1336]	[1899]	[3235]
Rule A: Stack-Experiment	6.80	6.54	6.66
	[1768]	[2028]	[3796]
Rule B: Treatment-Month	6.88	6.59	6.72
	[1768]	[2028]	[3796]

The table shows the mean ln[PM emissions (kg/month)] with the number of observations given in the brackets by different imputation rules in the control group, treatment group, and the whole sample. Observational unit is stack-month, excluding interregnum and mock trading periods, across 292 potential plants in the whole sample.

C.4. Market Replacement Rule for Missing CEMS Data

Subsection C.1 gives the data imputation rule for pollution we use for the purposes of our analysis. The goal of this rule is to estimate mean emissions as accurately as possible for plants that are missing some observations on pollution. In this subsection we show the data replacement rule that was used in real time by the market. This replacement rule has two purposes: filling gaps in the emissions record, *and* penalizing plants for non-reporting.

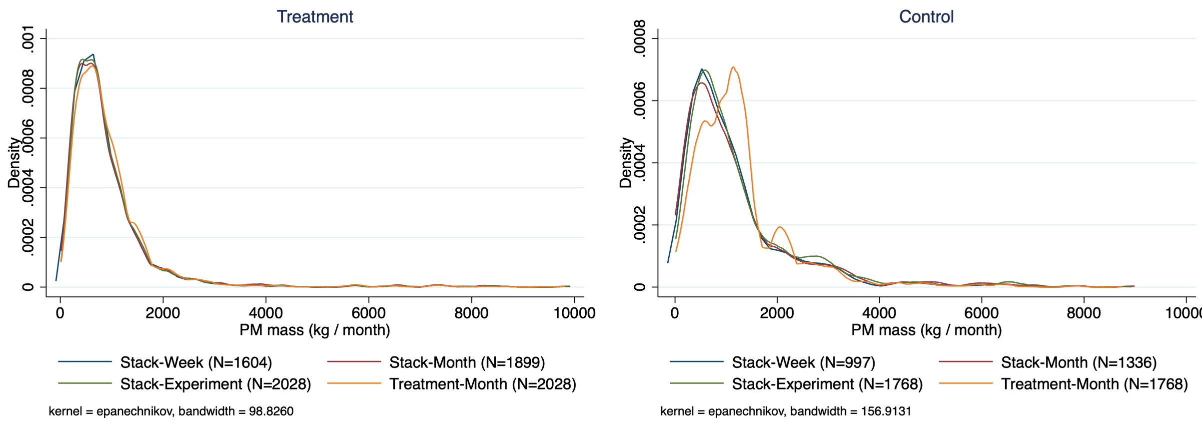
Table C2 shows the data replacement rule used in the market. The rule assumes that emissions are higher when data is missing for a longer period of time, in order to incentivize plants to report emissions reliably. By construction, the replacement rule used in the market will be upward biased relative to mean emissions during the time a plant is reporting.

C.5. Absence of Direct Effects of Monitoring on Emissions

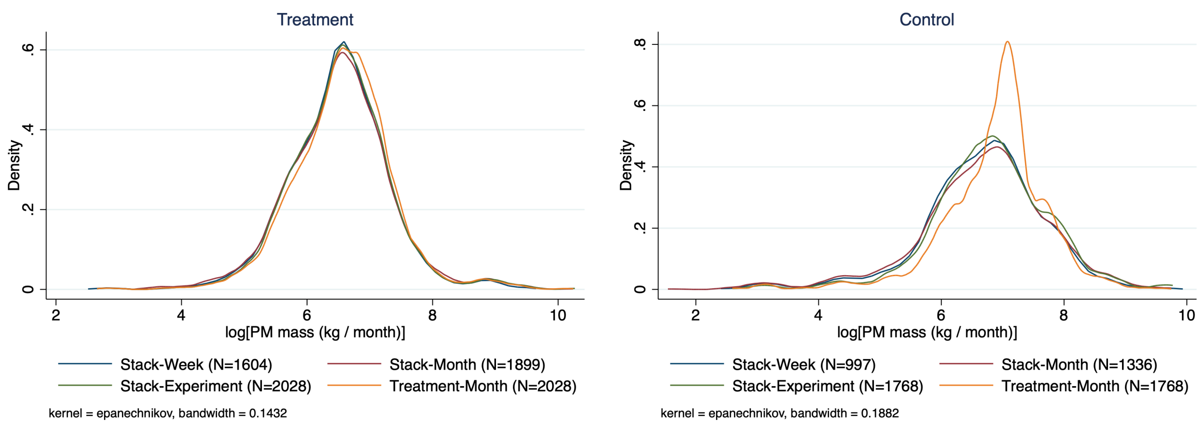
We have interpreted the control group in our evaluation as informative about outcomes under command-and-control regulation. One difference between the control and the status quo in Gujarat prior to the introduction of the market was that the control group also reports real-time emissions data using CEMS technology. This data underpins our experimental evaluation but could not be used to penalize plants since the legal notifications governing the status quo regime required that

Figure C4. Kernel density of PM emissions by treatment status

A. PM emissions



B. log(PM emissions)



This figure plots the kernel density of PM emissions (kg/month) in Panel A and log(PM emissions) in Panel B, both by treatment status, in different stages of imputation described in Section C.2. Stack-Week corresponds to the emissions variable after step 2. Stack-Month, Stack-Experiment, and Treatment-Month correspond to the variables constructed based on the No Imputation Rule, the Imputation Rule A, and the Imputation Rule B, respectively. Imputing the treatment group mean causes values to converge to the group mean, so the distribution of PM emissions and that of log(PM emissions) should have less dispersion under Rule B. Since the distribution of emissions is highly positive-skewed, the emissions of most plants are less than the group mean. Rule B, therefore, inflates the emissions of those plants. As a result, the peak of the kernel density curve under Treatment-Month for the control group shifts to the right. As the distribution of PM emissions is more clustered near the mean under Rule B, the mean of log(PM emissions) should be closer to the log of mean PM emissions for Rule B. By the concavity of log function, the log of mean is no less than the mean of log values. Hence, the mean of log(PM emissions) should be higher for Rule B than others.

regulatory action be based on pollution samples collected manually.

Here we ask whether CEMS monitoring, even without accompanying regulatory changes, might itself change plant behavior. We worked with GPCB to rollout CEMS as a randomized

Table C2. Imputation Rules for Missing CEMS Data

% Data Available During Week	Imputation for Missing Stack Data Values (kg/hr)
$>95\%$	Stack's own mean operating emissions load during the week
80-95%	Stack's own 75^{th} percentile emissions load during the week
50-80%	Stack's own 90^{th} percentile emissions load during the week
1-50%	Stack's own 90^{th} percentile emissions load during the three months prior to the start of the compliance period
$<1\%$	Flat rate of population emissions load (8.08 kg/hr)

The table gives the data replacement rule used in the emissions market. The left column shows the raw data availability during the week. The right column shows the imputation rule for each level of data availability.

experiment in order to test for such monitoring effects. Plants were randomly assigned to one of three phases. The random assignment means that plants receiving a late CEMS mandate form a valid control group for those with an early mandate.

We present results from a simple specification regressing measures of plant pollution obtained from manual measurements on CEMS treatment status. CEMS obviously cannot form the outcome measure for a treatment mandating CEMS installation. The pollution data comes from the result of manual emission samples carried out by the environmental regulator as part of their inspection schedule. We run the following regression:

$$y_{it} = \beta_0 + \beta_1 Treat_i \times Post_t + \alpha_i + \gamma_t + \varepsilon_{it}$$

where α_i is a plant fixed effect and γ_t is a month by year fixed effect. The dummy $Treat_i$ is 1 for plants in Phase 1 and 0 for plants in Phase 3. The outcome variable y is a measure of plant pollution from manual readings taken by the GPCB as part of their regular schedule of testing. The variable $Post_t$ is 0 for all control observations. For treatment (phase 1) units, it takes the value 1 once a plant has installed CEMS. β_1 is the treatment effect of CEMS on pollution.

Table C3 reports results from this regression. The main conclusion is that there is little evidence of differences in pollution between plants that had already installed CEMS relative to those that had not. Sudarshan (2023) provides related results including additional information on the rollout of

CEMS in Gujarat, a description of different technologies, and practical considerations associated with using this data for regulation.

Table C3. Effects of CEMS Installation on Plant Emissions

	PM, mg/Nm3 (1)	Log(PM) (2)
Treatment Effect	0.432	-0.0601
	(23.84)	(0.0912)
Observations	796	796
Year-Month FE	Yes	Yes
Plant fixed effects	Yes	Yes
Plants	197	197
R^2	0.3384	0.4276
Mean dependent variable	142.8	4.757

Dependent variable emission measures are from GPCB's regularly scheduled manual samplings. Unit of observation is plant. Sample is restricted to plants in Phases 1 and 3 of CEMS rollout. The treatment indicator in the regression is set to the interaction of the plant being in Phase 1 and not being in the control group. The treatment indicator is therefore set to 0 for the union of all experiment control plants and all Phase 3 plants. Standard errors are clustered at the plant-level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

C.6. Treatment effect on PM emissions with different imputation rules for the control and treatment groups

This section examines how treatment effects vary based on the stringency of the imputation rules applied to control and treatment plants. Table C4 presents the results. On the main diagonal, where the imputation rules are assumed the same, the treatment effect is as large or larger in magnitude as the preferred estimates. The estimated treatment effect grows larger with more punitive (higher quantile) imputation rules because control plants are missing more data than treatment plants. Thus increasing emissions for missing data increases control emissions more than treatment emissions and increases the magnitude of the estimated treatment effect.

Off the main diagonal below, the treatment imputation rule is assumed to be relatively higher than for the control. For example, in the first column, Rule A is maintained for imputation in the control group, roughly imputing emissions at the mean for the same plant during times when it is reporting. The rows give the estimated treatment effect if the imputation rule for the control group remains at Rule A but the imputation rule for the treatment group increases, corresponding to higher emissions assumptions in the treatment group when treatment CEMS are not reported. The treatment effect is similar to the main estimate, although somewhat reduced, even when missing emissions in the treatment group are imputed at the 80th percentile of treatment group emissions when reporting. The treatment effect is negative but not statistically significant if treatment group emissions are imputed at the 90th percentile of treatment group emissions when reporting, and close to zero if treatment group emissions are imputed at the market imputation rule, which fills in extremely high levels of emissions as punishment when a plant does not report.

Table C4. Treatment effect on PM emissions (log(PM mass/month)) with different imputation rules for the control and treatment groups

		Imputation rule – control				
		Rule A	p70	p80	p90	Market
		(1)	(2)	(3)	(4)	(5)
Imputation rule – treatment	(1) Rule A	-0.282***	-0.368***	-0.461***	-0.591***	-0.904***
		(0.074)	(0.076)	(0.076)	(0.077)	(0.072)
	(2) p70	-0.243***	-0.329***	-0.422***	-0.552***	-0.865***
		(0.074)	(0.075)	(0.076)	(0.077)	(0.072)
	(3) p80	-0.192**	-0.278***	-0.371***	-0.501***	-0.814***
		(0.074)	(0.075)	(0.076)	(0.077)	(0.072)
	(4) p90	-0.109	-0.196**	-0.288***	-0.418***	-0.731***
		(0.075)	(0.076)	(0.077)	(0.078)	(0.073)
	(5) Market	0.008	-0.078	-0.171**	-0.301***	-0.614***
		(0.076)	(0.077)	(0.078)	(0.079)	(0.074)

This table reports estimated treatment effects on PM emissions, as in Table III, column (5) of the main text, using different imputation rules for the treatment and control groups. The outcome variable is the log of plant-level PM mass (kg) per month. A detailed note on the construction of the outcome variable is in Appendix C.1. For each cell, the row describes the imputation rule used for treated plants and the column the imputation rule used for control plants. Rule A is stack-experiment imputation. Under this rule, missing values of a stack's daily PM mass rate are imputed using the stack's mean PM mass rate across the experiment (July 2019 to March 2021, excluding interregnum). p70 imputes missing values of a stack's daily PM mass rate using the stack's 70th percentile of PM mass rate across the experiment (July 2019 to March 2021, excluding interregnum). p80 and p90 are identical to the p70 imputation rule except that they use the 80th and 90th percentiles of PM mass rate respectively. Market is the market imputation rule described in Table C2. All regressions control for plant characteristics including capital expenditure, operating cost, log(total heat output), mean boiler installation year, and their corresponding indicators for missing values. Robust standard errors in parentheses are clustered at the plant level with statistical significance indicated by $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

D. APPENDIX: MODEL SPECIFICATION AND ABATEMENT COSTS

D.1. Model specification

1 *Abatement technology.* Plant i chooses the level of variable abatement expenditures Z_{it} in each compliance period $t = 1, 2, \ldots, 10$. Abatement expenditures could include running abatement equipment more, changing inputs like filters or chemicals more often, or devoting more labor to the maintenance and operation of a machine. Plants differ in total heat capacity H_i. Heat capacity is the steam production capacity of a boiler, analogous to the horsepower of a car engine, and is the relevant scale measure for fuel consumption and therefore air pollution emissions. Plants may also differ in other characteristics such as their abatement capital stock.

We let $Z_{it}(E_{it})$ be the level of expenditures as a function of emissions. Assume that $Z' < 0$ and $Z'' > 0$; expenditures are decreasing as a function of emissions but at a rate that decreases in magnitude as emissions grow. Further, there is some high, uncontrolled level of emissions $\overline{E}_i$ such that $Z_{it}(\overline{E}_i) = 0$. The plant spends an added fixed cost Z_{i0} to maintain its abatement capital. We treat this cost as sunk given the finding that abatement capital did not change in the experiment.

2 *Emissions market regulation.* An emissions market is a regulation that sets a market-level cap Q_t on emissions in period t and allows plants to trade permits so they collectively meet that limit. The regulator allocates permits A_{it} to each plant and may retain or sell the balance. In the Surat market, the allocation rule gave plants permits totaling 80% of the market cap in proportion to their heat capacity, $A_{it} \propto H_i$. Let P_t be the equilibrium price of permits, known to the plant.

Under the two assumptions of cost minimization and no market power, the plant chooses emissions to minimize the total cost of compliance:

$$\min_{E_{it}} Z_{i0} + Z_{it}(E_{it}) + P_t(E_{it} - A_{it}). \tag{6}$$

The first-order condition for the plant's problem under these assumptions is

$$-\frac{\partial Z_{it}(E_{it})}{\partial E_{it}} \;=\; MAC(E_{it}) = P_t. \tag{7}$$

This condition is the familiar one that the marginal abatement costs of the plant at the chosen emissions level equal the permit price. This equation has a unique solution for $E^*_{it} = MAC^{-1}(P_t)$ under our assumptions on the $Z(\cdot)$ function.

To calculate total costs we integrate marginal abatement costs to obtain each plant's total variable cost function. The marginal abatement cost function (4) we assume is consistent with a simple representation of total variable abatement costs:

$$Z_{it}(E_{it}) = e^{\beta_0+\tilde{\xi}_{it}} H^{\beta_2} \left(\frac{1}{\beta_1+1}\right) \left(\overline{E}_i^{\beta_1+1} - E_{it}^{\beta_1+1}\right), \qquad \beta_1 \in (-1,0), \tag{8}$$

where the parameters are common with (4). We estimate the parameters $\{\beta_1, \tilde{\xi}_{it}\}$ of the abatement cost function (8) using the marginal abatement cost specification (4). Moving from marginal to total variable abatement costs introduces a constant of integration. In (8), the constant $\overline{E}_i$ has a physical interpretation as the high level of uncontrolled emissions for a plant of size $H = 1$ when no variable abatement expenditures are made.

D.2. Calculating abatement costs by regime

We wish to compare abatement costs across regulatory regimes, but permit bids are only available in the treatment group. In this subsection we describe how we use the marginal abatement cost functions estimated in the treatment group to calculate abatement costs by regime.

1 *Abatement costs in the emissions market.* In the market all plants choose emissions to set their marginal abatement cost equal to the permit price, and therefore the marginal costs of all other plants. When all plants equalize their marginal abatement costs, the market as a whole reduces emissions at the lowest possible aggregate cost. The level of emissions depends on neither the plant's fixed costs of abatement Z_{i0} nor the initial permit allocation A_{it}.

Permit market equilibrium requires that aggregate emissions equal the market cap Q_t. Writing emissions as a function of the price, the equilibrium price is the P_t^* that solves

$$E_t(P_t^*) = \sum_i E_{it}(P_t^*) = Q_t. \tag{9}$$

The equilibrium price is unique because emissions for each plant monotonically decrease in price. At the equilibrium allocation, the total variable costs of abatement in the market can be written $Z_t^{ETS} = \sum_i Z_{it}(E_{it}^*)$, with plant emissions given by $E_{it}^* = E_{it}(P_t^*)$.

With our empirical specification of abatement costs we can solve for total variable abatement costs $Z_t^{ETS}(Q_t)$ at any proposed cap. The first step for a given cap Q_t is to solve (9) to find the equilibrium price P_t^*. With the estimated marginal abatement cost functions (4), the empirical inverse MAC function for each plant is:

$$E_{it}(P_t) \quad = \quad P_t^{1/\hat{\beta}_1} e^{-\hat{\xi}_{it}/\hat{\beta}_1}. \tag{10}$$

This function gives a plant's emissions as a function of the permit price. Substituting into (9), we can then find aggregate emissions at any price and solve for the equilibrium price $P_t^*(Q_t)$ for a given cap. We then calculate plant emissions with (10), evaluate plants' variable abatement costs (8) and sum across plants to find aggregate costs Z_t^{ETS}. The result of these steps is that we can write aggregate costs as a function of the aggregate emissions cap, $Z_t^{ETS}(Q_t)$.

2 *Abatement costs in the command-and-control regime* . We estimate the stringency of regulation in the command-and-control regime in the control group. A command-and-control regime is any rule that dictates emissions $\{E_{it}\}$ for each plant, rather than setting a limit across all plants. The current regime, *de jure*, sets a maximal concentration limit on pollution emissions. However, both in the control group and our prior work (Duflo et al., 2018), we observe *de facto* non-compliance with the intensity standard and fairly wide dispersion in emissions rates, rather than a point mass at the standard $\overline{R}$ (Online Appendix Figure B1).

We therefore estimate costs in the command-and-control regime by evaluating MAC functions at the observed emission rates in the control group. We represent emissions with plant-specific emissions rates $\overline{R}_{it} = E_{it}/H_i$ per unit of capacity. Since we observe emissions rates in the control group, it is straightforward to develop expressions for total emissions and total variable abatement costs. Total status quo emissions $E_t^{CC} = \sum_i H_i \overline{R}_{it}$ depend on the stringency of the plant-period specific intensity standards. Plant abatement costs are then the plant-period abatement cost function evaluated at this emissions level, $Z_{it}(H_i \overline{R}_{it})$. Summing across plants, total variable abatement costs under command-and-control are $Z_t^{CC} = \sum_i Z_{it}(H_i \overline{R}_{it})$.

In contrast to the outcome under an emissions market, there is no reason to expect that costs must be minimized by the command-and-control allocation of emissions. The de jure standard is a uniform concentration standard. There is widespread noncompliance even with this standard. We do not think this non-compliance is likely to equalize marginal abatement costs across plants. While our past work found that the regulator has some, albeit very noisy, information on pollution (Duflo et al., 2018), we expect marginal abatement costs are more difficult to estimate, since they cannot be observed directly on a plant visit. We therefore assume in our baseline case for the command-and-control regime that plant emissions rates are independent of plant marginal abatement costs.

We use five different representations of the status quo to capture the distribution of emissions rates in the command-and-control regime. The regimes differ in whether emissions rates are constant or dispersed across plants and whether they are conditioned on plant characteristics. The first two regimes we consider are: (i) *constant emissions rate* $R_{it} = \overline{R}$; (ii) *constant emissions rate with error* $\log R_{it} \sim \mathcal{N}(\mu_t, \sigma_t)$, fit separately in each period. These regimes are too simple to represent the status quo, because the data make clear that the emissions rate is declining in heat capacity. This fact is consistent with a regulatory regime that inspects large plants more often and so imposes greater expected penalties on them for high emissions rates.

We therefore favor regimes where the emissions rate depends on plant heat capacity. We fit the following regression in the control group separately for each compliance period:

$$\log R_{it} = \beta_{0t} + \beta_{1t} \log H_i + \varepsilon_{it}. \tag{11}$$

The remaining three regimes we consider follow this approach: (iii) *capacity-based emissions rate* $R_{it} = \exp(\widehat{logR_{it}})$; (iv) *capacity-based emissions rate with error* $R_{it} = \exp(\widehat{logR_{it}}\varepsilon_{it}^s)$ for draws $\varepsilon_{it}^s \perp \hat{\xi}_{it}$ from the residuals of (11); (v) *capacity-based emissions rate with correlated error*, similar to (iv), but with draws ε_{it}^s that are slightly negatively correlated ($\rho = -0.1$) with marginal abatement cost shocks $\hat{\xi}_{it}$.[33] We draw the emissions rate shocks from a log normal distribution fit to the variance of $\hat{\varepsilon}_{it}$ in each period. We include regime (iii) as a basis of comparison, though it will be biased due to the exponentiation of a predicted value fitted in logs.

We use these regimes to set counterfactual emissions rates, our proxy for intensity standards, for the treatment group plants, had they been regulated like control group plants. We then evaluate treatment plants' MAC functions at the simulated emissions rates to calculate the treatment plants' total abatement costs if they had been assigned to the control group.

In counterfactuals we evaluate costs not only at the distribution of emissions in the control group in the data, but also at higher or lower levels of emissions. We assume that a differently stringent command-and-control regime would scale up or down all emissions rates by a common factor δ. In the control group, we estimate fitted emissions rates across plants $\{\widehat{R}_{it}\}$ using one of the five regimes described above. We then calculate a scaling factor $\delta(Q_t) = Q_t/E_t^{CC}$ to meet emissions level Q_t. We evaluate plant-specific costs at alternate stringencies to calculate aggregate costs $Z_t^{CC}(\delta(Q_t)) = \sum_i Z_{it}(\delta(Q_t)H_i\widehat{R}_{it})$.

The idea of this approach is to preserve the dispersion in compliance, as observed in the current regime, while scaling emissions upwards or downwards to meet different possible caps. This assumes that the range of compliance at any new stringency would be the same, in proportional terms, as is observed in the control group. Since plant abatement costs are convex, this approach

[33]This implies that high-cost plants will have somewhat lower emissions rates. We introduce this correlation to capture, in a simple way, the observation that the regulator does have some information about plant emissions and targets more polluting plants more aggressively (Duflo et al., 2018).

of evaluating costs as we shift the distribution of emissions rates will produce higher aggregate abatement costs than would simply evaluating all plants at the new mean emissions rate.

3 *Comparison of abatement costs across regimes.* Online Appendix Table D1, panel A explores the robustness of the finding that the the market regime reduces total variable abatement costs to different approaches to assigning each plant's command-and-control emissions. We use the same two reference levels of aggregate emissions, 170 tons (columns 1 to 3) and 240 tons (columns 4 to 6). Row A reports equilibrium market price and total variable abatement costs under the market. Rows B1 to B5 present the total variable abatement costs under the command-and-control regime and its percentage difference, relative to costs in the emissions market. The rows under the command-and-control regime differ in how exactly they model the distribution of emissions.

There are two main findings. First, total variable abatement costs are lower under the emissions market than under the command-and-control regime. At the treatment emissions level, 170 tons per month, total variable abatement costs are 12% higher under the status quo (column 3, row B4) than under emissions trading (column 2, row A). The cost difference between regimes is great enough that costs are 6% lower under the emissions market—with a 30% cut in emissions—than in the command-and-control regime at the status quo emissions level (column 2, row A versus column 5, row B4).

Second, the cost differences among the alternative representations of the command and control regime are small and indeed smaller than the difference in cost between the market and command-and-control regimes. The differences in costs in the command and control regime are due to two forces: (i) heterogeneity in emissions rates interacting with convex abatement costs and (ii) scale effects.[34] We find that abatement costs are 8 to 13% higher under command-and-control at

[34]On heterogeneity, command and control regimes that allow idiosyncratic shocks across plants have higher costs than regimes that do not because abatement costs are convex. This convexity pushes up marginal abatement costs for plants that are assigned lower rates of emissions more than it reduces them for plants with higher rates (compare column 2, rows B1 and B2). On scale, we find that larger plants tend to have higher marginal abatement costs because the scale efficiencies in marginal abatement costs are outweighed by the higher marginal abatement costs associated with the more stringent emissions standards that large plants face (compare row B1 with rows B3 to B5).

the lower level of emissions (column 3) and 10 to 16% higher at the higher level of emissions. We prefer the simulation draws that condition on plant heat capacity and draw over residualized emissions levels, since emissions rates do vary systematically with plant size (heat capacity). For our preferred estimates, in row B4, the level of costs under the command-and-control regime are 12% and 15% higher, at the respective treatment and control levels of emissions, implying that the market cuts costs by 11% and 14% for these emissions levels.

Table D1. Variable abatement costs under alternative regulatory regimes

	Emissions = 170 tons			Emissions = 240 tons		
	Price (INR/kg) (1)	Cost (INR m) (2)	ΔCost (%) (3)	Price (INR/kg) (4)	Cost (INR m) (5)	ΔCost (%) (6)
Panel A: Iso-Elastic MAC Curve						
A. Emissions market	12.2	10.1	–	9.9	9.3	–
B. Command and Control						
1. Constant emissions rate		10.9	8.0%		10.2	10.0%
2. Constant emissions rate, with error		11.2	11.4%		10.6	14.1%
3. Capacity-based rate		10.9	8.2%		10.3	10.3%
4. Capacity-based rate, with error		11.3	11.8%		10.7	14.6%
5. Capacity-based rate, correlated error		11.4	12.9%		10.8	15.9%
Panel B: Step-Function MAC Curve						
A. Emissions market	15.3	6.0	–	11.6	5.0	–
B. Command and Control						
1. Constant emissions rate		7.9	32.4%		7.2	44.1%
2. Constant emissions rate, with error		7.9	32.6%		7.2	43.9%
3. Capacity-based rate		7.9	31.9%		7.1	42.9%
4. Capacity-based rate, with error		7.9	32.1%		7.1	42.9%
5. Capacity-based rate, correlated error		8.0	33.8%		7.3	45.5%

The table shows the results of counterfactual simulations under different regulatory regimes. Each row represents a different regime. Each panel corresponds to a different functional form assumption on the plant level marginal abatement cost curve. Within each panle the first row is the emissions market. The second through final rows in each panel are different command and control regimes that vary in how the emissions target is set for each plant. Constant emissions rate sets a single fixed ratio of emissions to heat output capacity for all plants. Constant emissions rate with error allows for idiosyncratic variation in the constant rate across plants. Capacity-based rate sets an emissions rate as a function of plant capacity, such that larger plants can have higher or lower rates of emission per unit capacity. Capacity-based rate with error allows for the capacity-based rate to idiosyncratically vary across plants. Finally, capacity-based rate with correlated error is the same as capacity-based rate with error except that the idiosyncratic error is drawn with a negative -0.1 correlation with estimated plant marginal abatement cost shocks. Columns 1 to 3 show results for emissions of 170 tons per month (the treatment level) and columns 4 to 6 for emissions of 240 tons per month (the control level). Within each set of three columns the variables show the market price (if applicable), the total variable abatement costs per month, and the change in abatement costs relative to the emissions market.

E. APPENDIX: MARGINAL COST CURVE SPECIFICATION

E.1. Uncontrolled Emissions

Without any abatement particulate emissions can be very high. We let $\overline{E}_i$ represent the uncontrolled level of emissions for plant i. To calculate it, we first find the average flow rate across manual samplings of that plant. These manual samplings include the ETS baseline, the CEMS baseline, as well as possible calibration measurements in 2020 and 2022. We then assume a maximum possible outlet concentration of 2500 $\frac{mg}{Nm^3}$, which was determined by our field staff to be a conservative plausible maximum, and we additionally assume operation for 12 hours per day for the entire period. These together determine a maximum possible emitted mass of PM for the stack for any period:

$$\overline{E}_i \frac{kg}{period} = \text{Flow-Rate}_i \frac{Nm^3}{hr} \cdot 12 \frac{hr}{day} \cdot 30 \frac{day}{period} \cdot 2500 \frac{mg}{Nm^3} \cdot \frac{1}{1,000,000} \frac{kg}{mg}$$

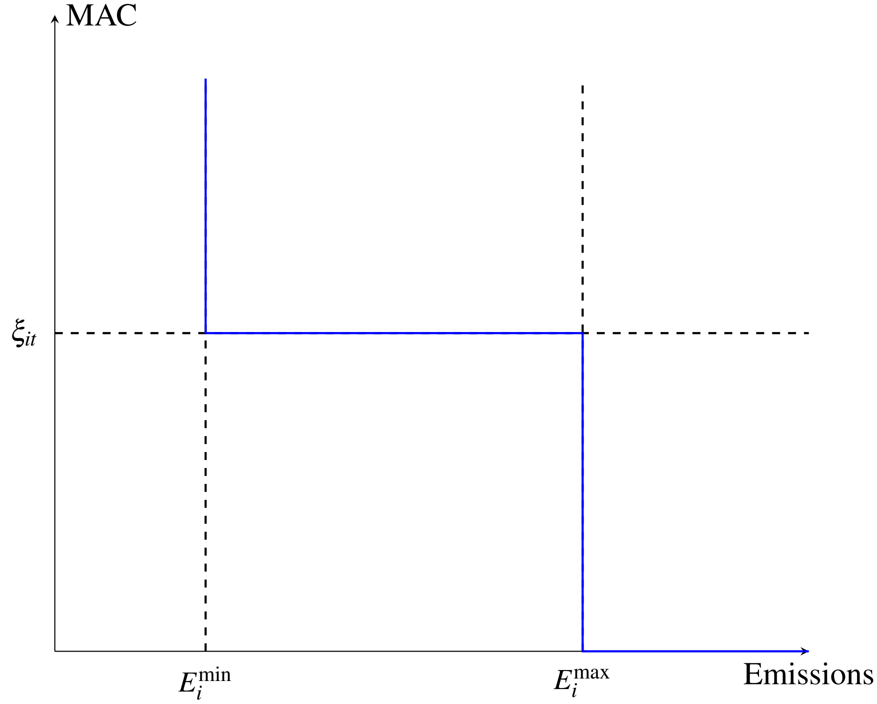
E.2. Step Function

In the body of the paper we assume that plants marginal abatement cost curves (hereafter MAC curves) follow a convenient iso-elastic form. However, it is worth examining what happens if we were to assume a different form of marginal abatement cost curve: We here consider a step-function MAC curve alternative.

1 *Functional Form.* The basic form of the step-function MAC curve, for a specific plant i in period t, is shown in Figure E1. The intuition for this functional form is that if plant i doesn't run their abatement technology they will emit at E_i^{max}, if they do run their abatement tech they will decrease their emissions down to E_i^{min}, and running their abatement tech costs ξ_{it}. In order to implement this MAC curve for each plant, we thus need, for each i,t, values of E_i^{min}, E_i^{max} and ξ_{it}.

We set $E_i^{\max} = \left(1 - \mathbf{1}\{i \text{ has cyclone}\} \cdot 0.8\right) \cdot \overline{E}_i$. See Appendix E.1 for how we set $\overline{E}_i$. This setup allows plants to run their cyclones "for free" when they have one, decreasing their max potential emissions by 80% (following the engineering estimates of cyclone efficiency reported in Table F1).

Figure E1. Example Step-Function MAC Curve for Plant *i* in Period *t*



Cyclones are mechanical abatement devices that run automatically based on the flow of the stack gas through the device. Thus, for those plants with a cyclone (which covers 98% of our sample) their abatement decision comes from whether or not they use further technology, with their cyclone being used at all times.

To calculate E_i^{min} we take the minimum of (a) the minimum observed level of emissions for plant *i* over any period, and (b) the abatement efficiency of plant *i*'s most advanced technology (again following efficiencies in Table F1) times their E_i^{max}. The intuition for this is that the emissions of a plant over a period provide an upper bound on how far they are able to abate their emissions, while the engineering estimates of the efficiency of their abatement tech represents a more direct estimate of how far they can reduce their emissions beyond E_i^{max}.[35]

Lastly, to calculate ξ_{it} we again assume that plants bid their marginal costs of abatement, and thereby set ξ_{it} to the corresponding fixed-effect of a regression of first half of period bid prices on

[35]For those plants which have a cyclone as their maximum abatement technology we set E_i^{min} to the minimum of (a) their minimum observed emissions over periods and (b) E_i^{max}, as the use of their cyclone has already been factored in to E_i^{max}.

plant-period fixed effects:

$$b_{itk} = \xi_{it} + \varepsilon_{itk}.$$

2 *Market Clearing Prices.* At a given market price P_t in period t, plant i will run their abatement equipment if and only if P_t is at least as high as their marginal abatement cost:

$$\text{Run}_{it}(P_t) = \mathbf{1}\{P_t \geq \xi_{it}\}$$

Therefore, E_{it} is a weakly decreasing function of P_t:

$$E_{it}(P_t) = E_i^{min} \cdot \text{Run}_{it}(P_t) + E_i^{max} \cdot (1 - \text{Run}_{it}(P_t))$$

We thus define the market clearing price in market period t, when there is market cap C_t as:

$$P_t^* = \min_{\mathbb{R}^+} P_t \text{ s.t. } \sum_i E_{it}(P_t) \leq C_t$$

The step functional form in which plants either don't abate at all or abate to their minimum means that it is possible for $\sum_i E_{it}(P_t^*) < C_t$. In this case, the discreteness of the abatement equipment leads to over-compliance.

At a given allocation of emissions across plants we use the model to calculate costs. The abatement cost function corresponds to the area under the MAC curve from E_{it} to ∞. However, the step functional form only allows E_{it} to be either E_i^{max} or E_i^{min} depending on Run_{it}. Abatement costs will then equal:

$$\begin{aligned} Z_{it}(E_{it}) &= \xi_{it} \cdot (E_i^{max} - E_{it}) \\ &= \text{Run}_{it} \cdot \xi_{it} \cdot (E_i^{max} - E_i^{min}) \\ &= \mathbf{1}\{E_{it} = E_i^{min}\} \cdot \xi_{it} \cdot (E_i^{max} - E_i^{min}) \end{aligned}$$

Command-and-Control As before we assume the command-and-control regime sets standards and associated emissions levels E_{it} for each plant-period. If this assigned $E_{it} < E_i^{min}$ then it will be impossible for plant i to abate to this level, so we re-assign $E_{it} = E_i^{min}$. We then assume that plants abate to achieve these levels of emissions, costing them:

$$Z_{it}(E_{it}) = \xi_{it} \cdot (E_i^{max} - E_{it})$$

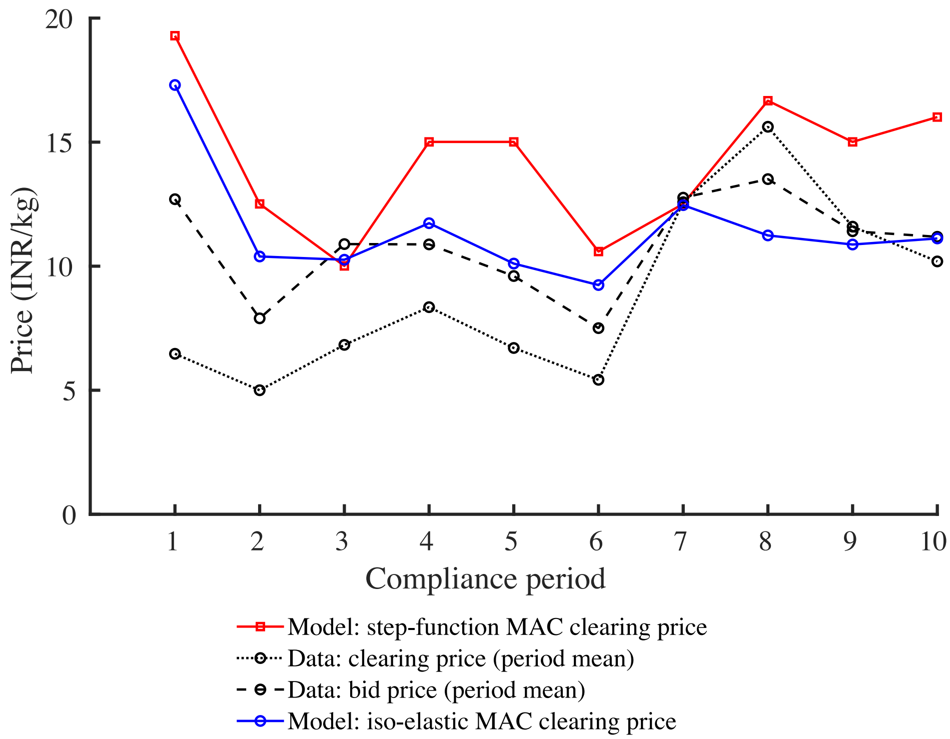
3 *Results.* We now duplicate several results from the paper using the alternative, step-function functional form for marginal abatement costs.

Figure E2 shows market prices calculated with the step-function form in red as a test of in-sample model fit. The step-function model tends to over-predict market prices, relative to the iso-elastic cost function model.

Figure E3 shows the distribution of emissions in the treatment market calculated with the step-function MAC (panels A and B), the iso-elastic MAC (panels C and D) and in the data. The step-function MAC leads to a multi-modal distribution of emissions. Because plants either run their equipment (if the cost of the step is low enough) or do not, emissions are dispersed and there are separate modes for plants based on these operating decisions. By contrast, the distributions of emissions for the iso-elastic MACs and in the data are smoother.

We carry the step-function results through to counterfactual analysis of market cost savings in Table D1, panel B. The rows of Panel B correspond to the different regimes discussed in 3 , simply changing the assumed functional form of the marginal abatement cost curves. The main finding is that the alternative, step-function MAC model predicts much larger counterfactual cost savings from the emissions market. The reason for this result is that, in the step function model, if a plant is estimated to have low marginal abatement costs it will always have low costs, up to the maximum efficacy of a piece of equipment. The costs of misallocation of abatement are therefore large because the model extrapolates in-sample differences in cost over a large range of emissions. In the iso-elastic MAC model, by contrast, a low-cost plant cannot take over such a large share of counterfactual abatement in the market, because its own MAC would curve upwards.

Figure E2. Model Fit to Market-Clearing Prices with Step-Function Alternative

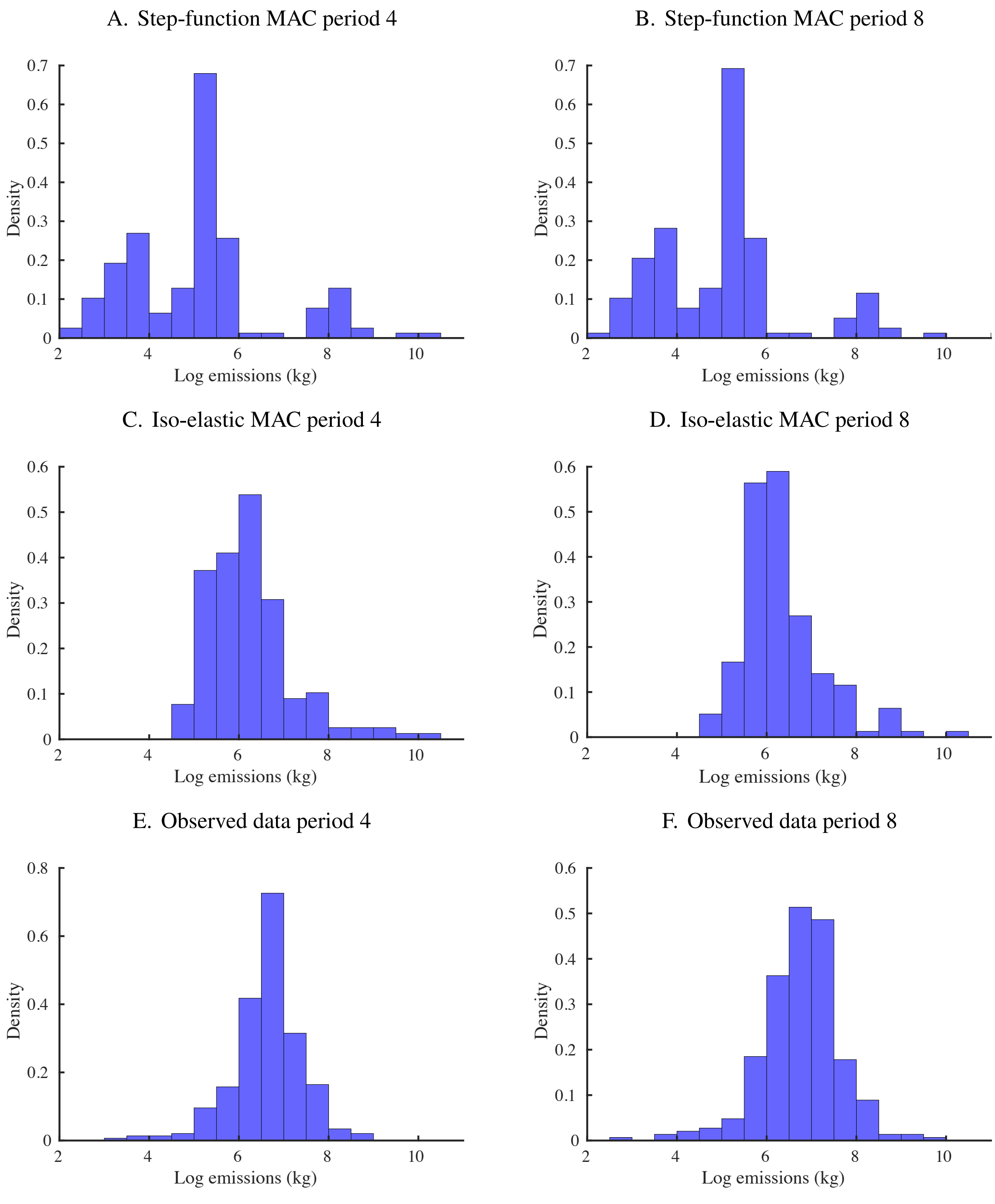


The figure shows the fit of the step-function and iso-elastic MAC models compared to the time series of market and bid prices by compliance period. The solid red line with square points is the time series of market-clearing prices in the fitted model with step-function MACs. The solid blue line with circular points is the time series of market-clearing prices in the fitted model with the original iso-elastic MACs. The models are fit based on bids in the first half of each compliance period. The dashed (black) line is the time series of mean bid prices in the data and the dotted (black) line is the time series of market-clearing prices.

E.3. Iso-elastic MAC with heterogeneous elasticities of abatement costs

Our main specification for marginal abatement cost curves (4) allows the mean log marginal abatement costs for each plant-period to differ but constrains all plants to have the same elasticity of MAC with respect to emissions. This part considers our counterfactual results if we allow heterogeneity in the MAC elasticity. In Table IV, column 5 allows the elasticity of MAC to differ by what abatement equipment a plant has installed. Table E1 replicates the counterfactual results of Table D1 Panel A, with the specification of column 5. We find that the magnitude and qualitative pattern of cost savings in the emissions trading regime relative to the control regime are similar to those reported in Table D1 Panel A.

Figure E3. Histograms of predicted versus observed emissions



The figure shows predicted and observed emissions levels in 2 periods for 2 different MAC curve specifications. Panels A and B show predicted emissions when running our emissions market model under a step-function MAC curve in periods 4 and 8 respectively. Panels C and D show the same except using the iso-elastic MAC curve. Panels E and F show the observed distribution of emissions in those periods.

Table E1. Variable abatement costs under alternative regulatory regimes (with Heterogeneity by APCD)

	Emissions = 170 tons			Emissions = 240 tons		
	Price (INR/kg) (1)	Cost (INR m) (2)	ΔCost (%) (3)	Price (INR/kg) (4)	Cost (INR m) (5)	ΔCost (%) (6)
A. Emissions market	12.2	10.1	–	10.0	9.3	–
B. Command and Control						
1. Constant emissions rate		11.0	8.6%		10.3	10.7%
2. Constant emissions rate, with error		11.3	11.8%		10.7	14.7%
3. Capacity-based rate		11.0	8.9%		10.4	11.1%
4. Capacity-based rate, with error		11.3	12.3%		10.8	15.2%
5. Capacity-based rate, correlated error		11.4	13.0%		10.8	16.2%

The table shows the results of counterfactual simulations under different regulatory regimes. Each row represents a different regime. The first row is the emissions market. The second through final rows are different command and control regimes that vary in how the emissions target is set for each plant. Constant emissions rate sets a single fixed ratio of emissions to heat output capacity for all plants. Constant emissions rate with error allows for idiosyncratic variation in the constant rate across plants. Capacity-based rate sets an emissions rate as a function of plant capacity, such that larger plants can have higher or lower rates of emission per unit capacity. Capacity-based rate with error allows for the capacity-based rate to idiosyncratically vary across plants. Finally, capacity-based rate with correlated error is the same as capacity-based rate with error except that the idiosyncratic error is drawn with a negative -0.1 correlation with estimated plant marginal abatement cost shocks. Columns 1 to 3 show results for emissions of 170 tons per month (the treatment level) and columns 4 to 6 for emissions of 240 tons per month (the control level). Within each set of three columns the variables show the market price (if applicable), the total variable abatement costs per month, and the change in abatement costs relative to the emissions market.

F. APPENDIX: ADDITIONAL RESULTS

F.1. Engineering Estimates of Abatement Costs

This section compares market prices for pollution permits to engineering estimates of the costs of running abatement equipment. In theory, the price of permits should reflect the marginal abatement costs to each plant. We check this assumption, in broad terms, by comparing permit prices to engineering measures of abatement costs.

To probe the validity of the assumption that bids can be used to infer marginal abatement costs, we compare the bids against engineering estimates of abatement costs from Indian air pollution control device vendors. As described in Section III, the market cleared at prices between the floor of INR 5 per kg and INR 15 per kg, though average bid prices ranged as high as INR 45 per kg. Online Appendix Table F1 presents estimates of abatement costs under *ideal* operating conditions for four kinds of air pollution control devices under four hypothetical plant configurations. This table assumes, as is likely the case in our data, that plants are already operating a single cyclone. Engineering abatement costs vary widely depending mainly on (i) the scale of the plant (ii) the type of equipment that is on the margin. If a plant is already running a cyclone, then average (marginal) abatement costs for a mid-size plant (6 ton per hour boiler) to operate an additional cyclone are 7 (2) INR per kg and an additional bag filter 10 (3) INR per kg. If a plant is small and already running a cyclone, average (marginal) abatement costs to run a dry scrubber are as high as 71 (21) INR per kg. Variable abatement costs therefore range from INR 2 per kg to INR 20 per kg, depending on what piece of equipment is used, under the assumed, ideal operating efficiency. If operating efficiency is actually lower, as seems likely, and the reduction in emissions therefore smaller, then the abatement cost per kg of emissions reduction would increase inversely with the decline in efficiency.

Overall, this exercise supports the assumption that the bidding data can be used to infer marginal abatement costs. We find that the market clearing permit prices overlap with engineering estimates of the marginal abatement costs associated with operating abatement equipment.

Table F1. Engineering estimates of abatement costs under ideal operating efficiency, if a cyclone is already operating

	Cyclone (1)	Bag Filter (2)	Scrubber (3)	ESP (4)
	Total Boiler Capacity = 3 TPH			
Capital costs (Rs/month, amort.)	6953.33	6518.75	10430.00	78225.00
Variable costs (Rs/month)	3000.00	2812.50	4500.00	33750.00
Emission reduction (%)	80.00	99.00	94.00	99.70
Assumed pollution (kg/month)	1575.90	1575.90	1575.90	1575.90
Emission abatement (kg/month)	1260.72	1560.14	1481.34	1571.17
Average abatement cost (Rs/kg)	7.89	5.98	10.08	71.27
Variable abatement cost (Rs/kg)	2.38	1.80	3.04	21.48
	Total Boiler Capacity = 6 TPH			
Capital costs (Rs/month, amort.)	9560.83	15645.00	16514.17	104300.00
Variable costs (Rs/month)	4125.00	6750.00	7125.00	45000.00
Emission reduction (%)	80.00	99.00	94.00	99.70
Assumed pollution (kg/month)	2323.37	2323.37	2323.37	2323.37
Emission abatement (kg/month)	1858.70	2300.14	2183.97	2316.40
Average abatement cost (Rs/kg)	7.36	9.74	10.82	64.45
Variable abatement cost (Rs/kg)	2.22	2.93	3.26	19.43
	Total Boiler Capacity = 8 TPH			
Capital costs (Rs/month, amort.)	11299.17	19990.83	26075.00	173833.33
Variable costs (Rs/month)	4875.00	8625.00	11250.00	75000.00
Emission reduction (%)	80.00	99.00	94.00	99.70
Assumed pollution (kg/month)	3612.38	3612.38	3612.38	3612.38
Emission abatement (kg/month)	2889.91	3576.26	3395.64	3601.55
Average abatement cost (Rs/kg)	5.60	8.00	10.99	69.09
Variable abatement cost (Rs/kg)	1.69	2.41	3.31	20.82
	Total Boiler Capacity = 15 TPH			
Capital costs (Rs/month, amort.)	13906.67	20860.00	26075.00	234675.00
Variable costs (Rs/month)	6000.00	9000.00	11250.00	101250.01
Emission reduction (%)	80.00	99.00	94.00	99.70
Assumed pollution (kg/month)	8781.49	8781.49	8781.49	8781.49
Emission abatement (kg/month)	7025.19	8693.67	8254.60	8755.14
Average abatement cost (Rs/kg)	2.83	3.43	4.52	38.37
Variable abatement cost (Rs/kg)	0.85	1.04	1.36	11.56

Note. Table displays engineering estimates of abatement cost for different APCDs and boiler capacities. We assume one cyclone is already operating when calculating the quantity of abatement, and we assume each APCD is purchased in isolation. Costs can be compared with those in other tables at a rate of INR 70 to USD 1. Capital costs are amortized to a monthly flow value. All plants are assumed to have a raw inlet concentration of 2,000 mg/Nm^3; in practice it can vary between 1,000 mg/Nm^3 and 10,000 mg/Nm^3. This is converted to a monthly mass rate via a volumetric flow rate collected at baseline, assuming continuous operation for 16 hours/day and 25 days/month. Of plants with boilers in our analysis sample, the boiler capacity (BC) distribution is: 11% have 2-3 TPH BC, 47% have 4-7 TPH BC, 36% have 8-14 TPH BC, 6% have 15+ TPH BC.

F.2. Treatment effects on capital installation

Table F2 shows that the ETS treatment is estimated to have no effect on the presence of air pollution control devices (APCDs), overall, since all plants already have APCDs of some kind installed. There is suggestive evidence of a small shift toward less expensive APCDs such as cyclones and bag filters (columns 1 and 2).

Table F2. Treatment effects on the presence of abatement devices

	All APCDs	Components			
		Cyclone	Bag	Scrubber	ESP
	(1)	(2)	(3)	(4)	(5)
ETS Treatment (=1)	0	0.0233*	0.0650***	-0.0151	-0.0311
	(.)	(0.0134)	(0.0231)	(0.0310)	(0.0207)
R^2	.	0.66	0.68	0.71	0.75
Control mean	1.00	0.95	0.85	0.67	0.12
Plants	276	276	276	276	276

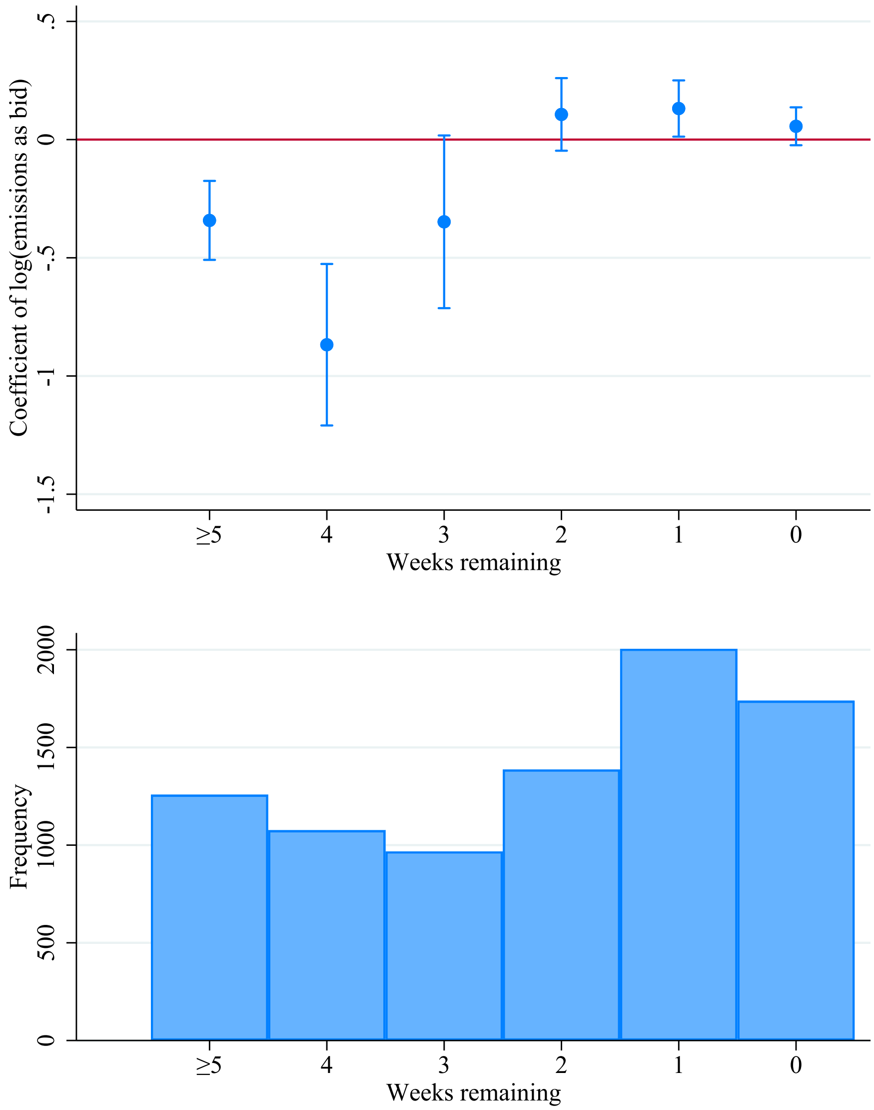
This table reports the effects of treatment assignment on the presence of APCDs. All specifications control for the corresponding baseline value. Robust standard errors are given in parentheses with statistical significance indicated by $^{*}p<0.10$; $^{**}p<0.05$; $^{***}p<0.01$.

F.3. Model robustness checks

1 *Heterogeneity in estimated elasticities by time of bid.* Section V estimates the elasticity of marginal abatement costs with respect to emissions using data from the first half of each compliance period. The argument is that plants only have a choice between abatement and the purchase of permits during the first half of the period, because by the end of a period, emissions are sunk and plant willingness-to-pay for permits should not depend on their abatement costs.

Figure F1 tests this idea by estimating the same elasticity separately in each week of the compliance period. We find that the elasticity of marginal abatement costs with respect to emissions is negative and economically and statistically significant during the first several weeks of the compliance period. When there are two weeks or less remaining in the compliance, by contrast, the same elasticity is estimated to be close to zero. As expected, plants' bids are not sensitive to abatement costs when there is little time left in a period in which to abate.

Figure F1. Elasticity estimate by weeks remaining in the order period



The top panel presents the coefficients of log(emissions as bid) from regressing ln(bid price) on log(emissions as bid) and plant × period fixed effects, estimated with different sample truncations defined by the number of weeks remaining in the order period. The bottom panel shows the number of bids placed in different sample truncations.

F.4. Significance of Emissions Non-Reporting

In this section we investigate if differences between plants who report different amounts might be driving any of our results.

Table F3 shows the same balance table as Table I from the main paper, only now comparing those plants above versus below median levels of daily data reporting. There are some differences between high- and low-reporting plants, but they are not large. The main difference appears to be that high-reporting plants are somewhat larger, with higher sales revenue and boiler house capital expenditure (panel A). There are no differences in abatement equipment installation (panel B). On emissions, high-reporting plants have similar PM emissions mass rates and baseline PM concentrations to low-reporting plants, and high-reporting plants are somewhat more likely (10 pp with a standard error of 5.5 pp) to be above the emissions concentration standard at baseline.

We next examine whether abatement costs or being in the treatment group are related to plant's data reporting. To do this we calculate predicted abatement costs using baseline rates of equipment installation for different plants and the average costs from Table F1. We then regress reporting rates on predicted abatement costs and their interaction with treatment status to test whether plants with higher costs report less. The results of this regression are in Table F4. We do not find any significant effect of predicted abatement costs on reporting or differential reporting in the treatment.

Table F5 then estimates treatment effects on pollution controlling directly for plant reporting rates. We estimate a similar magnitude of average treatment effect conditional on reporting rates.

Table F3. Balance of plant characteristics by whether report more than median reporting

	Over Median (1)	Under Median (2)	Difference (3)
Panel A: Plant Measures			
Total electricity cost (1,000 USD)	466.5	344.1	122.4
	[833.7]	[401.9]	(78.1)
Log(plant total heat output)	15.6	15.6	0.042
	[0.62]	[0.48]	(0.065)
Size as recorded on environment consent (1 to 3)	1.41	1.31	0.10
	[0.66]	[0.59]	(0.074)

Small-scale (size=1)	0.68	0.75	-0.068
	[0.47]	[0.43]	(0.054)
Large-scale (size=3)	0.096	0.062	0.033
	[0.29]	[0.24]	(0.032)
Number of stacks	1.08	1.05	0.030
	[0.40]	[0.21]	(0.037)
Textiles sector (=1)	0.85	0.88	-0.028
	[0.36]	[0.33]	(0.041)
Gross Sales Revenue in 2017 (1,000 USD)	15125.8	6950.0	8175.8*
	[54715.6]	[13111.3]	(4609.6)
Panel B: Plant Abatement and Investment Cost			
Boiler house employment	36.4	32.7	3.78
	[32.4]	[30.0]	(3.70)
Boiler house capital expenditure (1,000 USD)	215.8	149.6	66.2*
	[403.6]	[174.7]	(36.8)
Boiler house operating cost (1,000 USD)	142.8	108.1	34.7*
	[202.5]	[83.9]	(17.8)
APCD: Cyclone present	0.98	0.97	0.012
	[0.14]	[0.17]	(0.019)
APCD: Bag filter present	0.82	0.86	-0.038
	[0.38]	[0.35]	(0.043)
APCD: Scrubber present	0.64	0.62	0.020
	[0.48]	[0.49]	(0.058)
APCD: ESP present	0.12	0.070	0.051
	[0.33]	[0.26]	(0.035)
Panel C: Plant Pollution Measures			
Plant total PM mass rate (kg/hr)	3.70	3.50	0.20
	[4.97]	[3.73]	(0.52)
Plant mean PM concentration (mg/Nm^3)	179.5	167.8	11.7
	[154.2]	[152.3]	(18.2)
Plant mean Ringelmann score (1 to 5)	1.32	1.41	-0.086*
	[0.39]	[0.41]	(0.047)
Above regulatory standard at ETS baseline (=1)	0.36	0.26	0.10*
	[0.48]	[0.44]	(0.055)
Number of plants	156	136	

This table shows differences in plant scale (panel A), plant abatement and investment costs (panel B), and plant pollution (panel C) between the plants who report above and below the median level across plants. Each plants level of reporting is calculated as the average minute-level CEMS data availability across the full sample period and across all stacks belonging to that plant. The only plants which are included in this table are those in the analysis set. This sample consists of 292 plants that had at least one day of PM data from CEMS devices during the ETS experiment. In panel B, cyclone, bag filter, scrubber, and electrostatic precipitator (ESP) are different air pollution control devices (APCDs). Some plants did not respond to some questions in the survey and so certain variable rows have fewer observations than the full sample size. The first and second columns show means with standard deviations given in brackets. The third column shows the coefficients from regressions of each variable on treatment, with robust standard errors in parentheses. $^{*}p<0.10$; $^{**}p<0.05$; $^{***}p<0.01$.

Table F4. Treatment effect on reporting by predicted plant abatement costs

	Dependent Variable: Share of Day Not-Reporting
Treatment (=1)	-0.166***
	(0.035)
Predicted Abatement Cost	0.000
	(0.002)
Predicted Abatement Cost × Treatment	0.000
	(0.002)
R^2	0.13
Observations	304

Unit of observation is plant. Predicted Abatement Cost variable for industry set to the engineering estimate of the average abatement cost per kg from Table F1 assuming Boiler Capacity = 8TPH) for the most advanced abatement technology of the plant: Cyclone less advanced than bag-filter less advanced than scrubber less advanced than ESP. Share of day not reporting calculated at industry level as average over industry's stacks and over all days (excluding interregnum). Robust standard errors in parentheses. Statistical significance is indicated by $^{*}p<0.10$; $^{**}p<0.05$; $^{***}p<0.01$.

Table F5. Treatment effects on PM emissions (log(PM mass/month)) controlling for data availability

	No Imputed Months				Imputed Months			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ETS Treatment (=1)	-0.179**	-0.203***	-0.180**	-0.207***	-0.240***	-0.247***	-0.233***	-0.250***
	(0.077)	(0.077)	(0.076)	(0.076)	(0.073)	(0.073)	(0.061)	(0.061)
Share of Day Reporting	0.0000343	0.00127	0.000325	0.00148	-0.00222***	-0.00187**	-0.00445***	-0.00350***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Year-Month FE		Yes		Yes		Yes		Yes
Imputation rule					Rule A	Rule A	Rule B	Rule B
Reweighted			Yes	Yes				
Mean dep. var (control)	6.67	6.67	6.66	6.66	6.80	6.80	6.88	6.88
R^2	0.13	0.18	0.14	0.17	0.19	0.22	0.19	0.27
Plants	292	292	292	292	292	292	292	292
Observations	3235	3235	3235	3235	3796	3796	3796	3796

This table reports the estimated treatment effects on PM emissions adding average availability. The outcome variable is the log of plant-level PM mass (kg) per month. A detailed note on the construction of the outcome variable is in Appendix C.1. Columns 5 and 6 impute data with Imputation Rule A: *Stack-Experiment*. Under this rule, missing values of a stack's daily PM mass rate are imputed using the stack's mean PM mass rate across the experiment (July 2019 to March 2021, excluding interregnum). Columns 7 and 8 impute data with Imputation Rule B: *Treatment-Month*. Under this rule, missing values of a stack's daily PM mass rate are imputed using the monthly mean PM mass rate of the stack's treatment group. All columns control for plant characteristics including capital expenditure, operating cost, log(total heat output), mean boiler installation year, and their corresponding indicators for missing values. In addition to plant controls, columns 2, 4, 6, and 8 add year-month fixed effects to control for time variant differences common in each plant. We also apply the inverse probability weighting method in columns 3 and 4. The probability of reporting in a month is predicted using a probit model where the only explanatory variable is an indicator variable for the treatment status in a prior experiment that randomized CEMS installation timing. Robust standard errors in parentheses are clustered at the plant level with statistical significance indicated by $^{*}p<0.10$; $^{**}p<0.05$; $^{***}p<0.01$.

F.5. Impact of COVID-19 Pandemic

In this section we include several analyses trying to determine the impact of COVID-19 on the experiment.

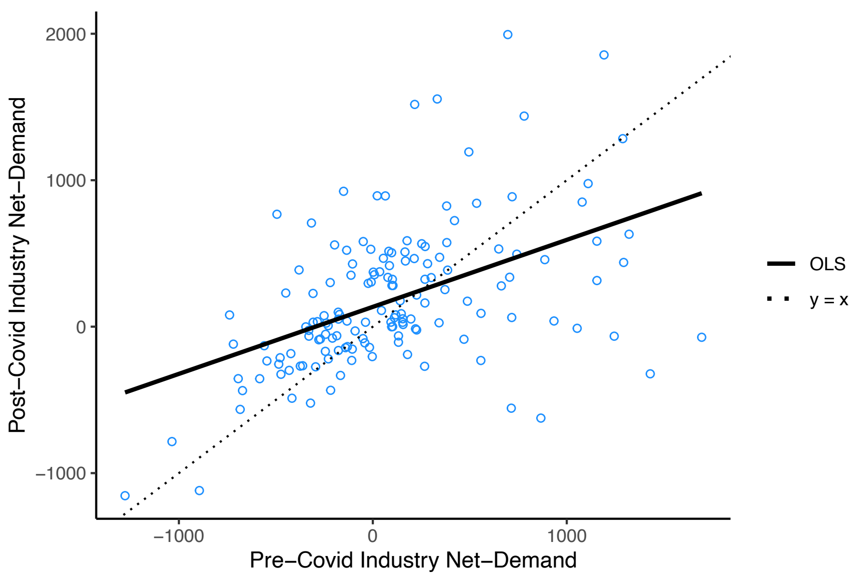
First, we examine how net-demand (defined as a plant's total period emissions less their initial permit allocation) differed before and after the COVID interregnum. Figure F2 shows a scatter plot of the plant net permit demand before (compliance periods 1 to 6) and after (periods 7 to 10) the COVID-19 interruption. The scatter plot shows that plant net demands are highly correlated in the pre- and post-Covid periods. Plants that have higher emissions than permit allocations before the pandemic tend to also have higher emissions than permit allocations afterwards.

In Table F6 we then estimate the treatment effect on emissions separately for the pre- and post-COVID subsets of our sample. In specifications without imputation or with imputation Rule A there is no statistically significant difference in the treatment effect before and after the Covid-19 lockdown. In specifications with imputation rule B, the treatment effect is statistically smaller in magnitude (less negative) after the lockdown but remains large, negative and statistically significantly different from zero. The point estimates for the treatment effect on pollution are smaller post-lockdown, which would be consistent with a less tightly binding cap in a weaker economy.

In Table F7 we re-estimate the counterfactual market versus command-control results from Table D1 using both the iso-elastic and step-function MAC, only restricting the sample to the pre-COVID periods.

Lastly, Figure F3 displays emissions over final permit holdings without GPCB's period 7 permit adjustment. This figure is identical to Figure IV other than that it removes additional permits granted to plants during compliance period 7, the initial post-Covid-lockdown period. Footnote 22 describes the adjustment in more detail.

Figure F2. Net demand before and after COVID



The figure shows net demand (emissions – initial allocation) for each industry averaged across pre vs post COVID compliance periods. Each point represents a single industry. The solid black line is the OLS fit for the data The dotted black line is the $y = x$ line. We omit the 9 industries with values of magnitude greater than 2000 for ease of visualization.

Table F6. Treatment effects on PM emissions (log(PM mass/month)) before and after COVID

	No Imputation				With Imputation			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ETS Treatment (=1)	-0.211**	-0.219**	-0.211**	-0.224***	-0.303***	-0.303***	-0.372***	-0.372***
	(0.087)	(0.088)	(0.085)	(0.085)	(0.077)	(0.077)	(0.060)	(0.060)
Post-Covid (=1)	-0.173**		-0.159**		-0.149***		-0.276***	
	(0.082)		(0.079)		(0.047)		(0.059)	
Treatment × Post-Covid	0.0652	0.0656	0.0668	0.0717	0.0538	0.0538	0.144**	0.144**
	(0.093)	(0.094)	(0.091)	(0.092)	(0.060)	(0.060)	(0.071)	(0.071)
Year-Month FE		Yes		Yes		Yes		Yes
Imputation rule					Rule A	Rule A	Rule B	Rule B
Reweighted			Yes	Yes				
Mean dep. var (control)	6.67	6.67	6.66	6.66	6.80	6.80	6.88	6.88
R^2	0.14	0.17	0.14	0.17	0.19	0.22	0.18	0.26
Plants	292	292	292	292	292	292	292	292
Observations	3235	3235	3235	3235	3796	3796	3796	3796

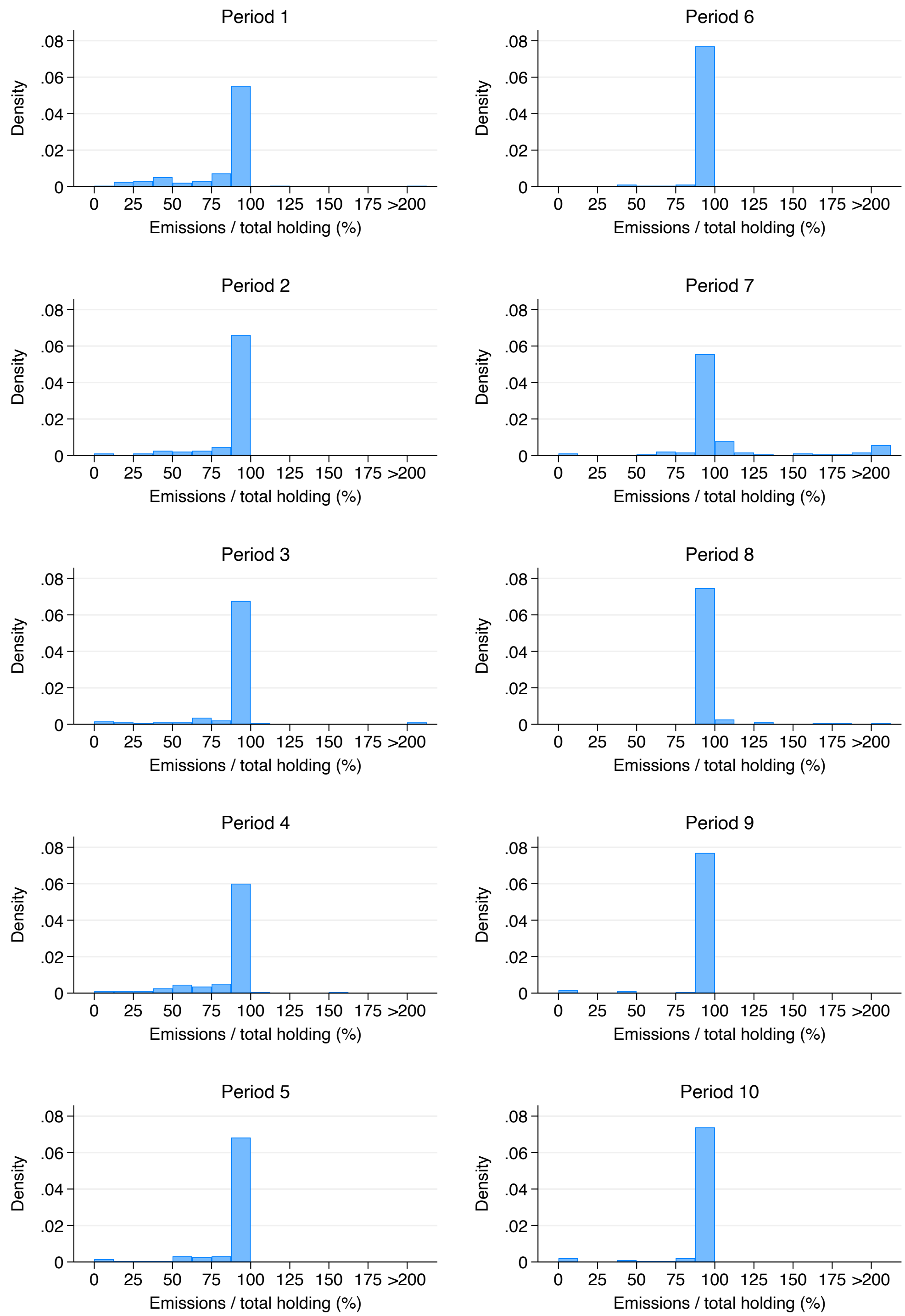
This table reports the estimated treatment effects on PM emissions. The outcome variable is the log of plant-level PM mass (kg) per month. A detailed note on the construction of the outcome variable is in Appendix C.1. Post-Covid is defined as periods 7 to 10. Columns 5 and 6 impute data with Imputation Rule A: *Stack-Experiment*. Under this rule, missing values of a stack's daily PM mass rate are imputed using the stack's mean PM mass rate across the experiment (July 2019 to March 2021, excluding interregnum). Columns 7 and 8 impute data with Imputation Rule B: *Treatment-Month*. Under this rule, missing values of a stack's daily PM mass rate are imputed using the monthly mean PM mass rate of the stack's treatment group. All columns control for plant characteristics including capital expenditure, operating cost, log(total heat output), mean boiler installation year, and their corresponding indicators for missing values. In addition to plant controls, columns 2, 4, 6, and 8 add year-month fixed effects to control for time variant differences common in each plant. We also apply the inverse probability weighting method in columns 3 and 4. The probability of reporting in a month is predicted using a probit model where the only explanatory variable is an indicator variable for the treatment status in a prior experiment that randomized CEMS installation timing. Robust standard errors in parentheses are clustered at the plant level with statistical significance indicated by $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table F7. Variable abatement costs under alternative regulatory regimes using only pre-COVID data

	Emissions = 170 tons			Emissions = 240 tons		
	Price (INR/kg) (1)	Cost (INR m) (2)	ΔCost (%) (3)	Price (INR/kg) (4)	Cost (INR m) (5)	ΔCost (%) (6)
Panel A: Iso-Elastic MAC Curve						
A. Emissions market	12.7	8.1	–	9.7	7.3	–
B. Command and Control						
1. Constant emissions rate		9.0	11.4%		8.3	13.6%
2. Constant emissions rate, with error		9.4	16.9%		8.8	20.4%
3. Capacity-based rate		8.9	10.9%		8.2	12.9%
4. Capacity-based rate, with error		9.4	16.7%		8.7	19.9%
5. Capacity-based rate, correlated error		9.5	18.2%		8.9	21.8%
Panel B: Step-Function MAC Curve						
A. Emissions market	16.3	5.8	–	12.0	4.8	–
B. Command and Control						
1. Constant emissions rate		9.0	54.6%		8.2	72.7%
2. Constant emissions rate, with error		9.0	54.7%		8.2	72.3%
3. Capacity-based rate		8.9	52.5%		8.0	68.7%
4. Capacity-based rate, with error		8.9	52.7%		8.0	68.5%
5. Capacity-based rate, correlated error		9.0	54.7%		8.2	71.8%

The table shows the results of counterfactual simulations under different regulatory regimes. Each row represents a different regime. The first row is the emissions market. The second through final rows are different command and control regimes that vary in how the emissions target is set for each plant. Constant emissions rate sets a single fixed ratio of emissions to heat output capacity for all plants. Constant emissions rate with error allows for idiosyncratic variation in the constant rate across plants. Capacity-based rate sets an emissions rate as a function of plant capacity, such that larger plants can have higher or lower rates of emission per unit capacity. Capacity-based rate with error allows for the capacity-based rate to idiosyncratically vary across plants. Finally, capacity-based rate with correlated error is the same as capacity-based rate with error except that the idiosyncratic error is drawn with a negative -0.1 correlation with estimated plant marginal abatement cost shocks. Columns 1 to 3 show results for emissions of 170 tons per month (the treatment level) and columns 4 to 6 for emissions of 240 tons per month (the control level). Within each set of three columns the variables show the market price (if applicable), the total variable abatement costs per month, and the change in abatement costs relative to the emissions market. Data used for estimation is restricted to pre-COVID periods (periods 1 to 6) only.

Figure F3. Distribution of Emissions over Final Permit Holdings by Compliance Period without GPCB's Period 7 Adjustment

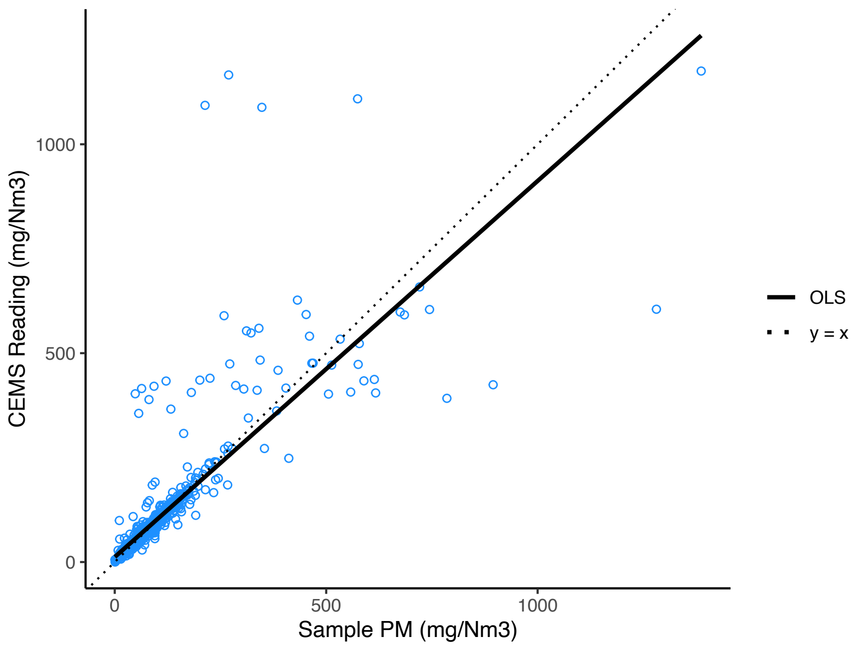


This figure plots the distributions of (emissions / final permit holdings $\times$ 100%) across treated plants ($N = 156$) by compliance period. Final permit holdings are the total number of permits a plant held at the end of the true-up period after each compliance period. Emissions data and permit holdings are from the administrative records of the market operator. Permit holdings are adjusted to remove those granted in GPCB's period 7 adjustment. Emissions are the validated emissions for each plant, which include any imputed emissions filled-in for periods of missing data. These validated emissions are used to determine compliance.

F.6. Manual Sampling and CEMS Comparison

Figure F4 plots emissions measurements from manual samplings versus CEMS readings of the emissions from the same window of time during which the manual sampling was taking place. A regression line and the $y = x$ line are also shown. There is a high correlation between the manual samples and CEMS readings.

Figure F4. Simulatenous CEMS and sampling comparison



The figure plots CEMS readings against concurrent manual samplings. Unit of observation is an industry. The solid black line is the OLS fit for the data The dotted black line is the $y = x$ line. We restrict the graph to only those CEMS readings with at least 15% data availability during the appropriate window, and to those with concentrations less than 2000 mg/Nm3.

F.7. Impact of Device Types

In Table F8 we estimate the treatment effect on emissions levels separately for each potential CEMS device type (type 1, types 2, or both types). We do not find any significant effects of device type on treatment effect. The coefficient on Type 2 devices is positive, though statistically insignificant. A positive effect would be consistent with larger, more sophisticated plants selecting a more expensive piece of equipment.

Table F8. Treatment effects on PM emissions (log(PM mass/month)) depending on device type

	No Imputation				With Imputation			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ETS Treatment (=1)	-0.130	-0.147*	-0.135	-0.154*	-0.247***	-0.247***	-0.310***	-0.310***
	(0.087)	(0.087)	(0.085)	(0.085)	(0.084)	(0.084)	(0.065)	(0.065)
Device FE: Type 2	0.159	0.177	0.141	0.159	0.0955	0.0955	0.126	0.126
	(0.219)	(0.221)	(0.221)	(0.224)	(0.173)	(0.173)	(0.136)	(0.137)
Device FE: Both Types	0.362*	0.354*	0.362*	0.355*	0.289	0.289	0.179	0.179
	(0.212)	(0.209)	(0.204)	(0.201)	(0.210)	(0.211)	(0.156)	(0.156)
Treatment × Device FE: Type 2	-0.0910	-0.0934	-0.0639	-0.0652	-0.00556	-0.00556	-0.0495	-0.0495
	(0.264)	(0.266)	(0.267)	(0.271)	(0.228)	(0.228)	(0.190)	(0.190)
Treatment × Device FE: Both Types	-0.319	-0.302	-0.311	-0.299	-0.229	-0.229	-0.0798	-0.0798
	(0.269)	(0.268)	(0.266)	(0.265)	(0.261)	(0.262)	(0.214)	(0.214)
Year-Month FE		Yes		Yes		Yes		Yes
Imputation rule					Rule A	Rule A	Rule B	Rule B
Reweighted			Yes	Yes				
Mean dep. var (control)	6.65	6.65	6.64	6.64	6.76	6.76	6.87	6.87
R^2	0.14	0.18	0.14	0.18	0.18	0.22	0.17	0.26
Plants	279	279	279	279	279	279	279	279
Observations	3110	3110	3110	3110	3627	3627	3627	3627

This table reports the estimated treatment effects on PM emissions. The outcome variable is the log of plant-level PM mass (kg) per month. A detailed note on the construction of the outcome variable is in Appendix C.1. Columns 5 and 6 impute data with Imputation Rule A: *Stack-Experiment*. Under this rule, missing values of a stack's daily PM mass rate are imputed using the stack's mean PM mass rate across the experiment (July 2019 to March 2021, excluding interregnum). Columns 7 and 8 impute data with Imputation Rule B: *Treatment-Month*. Under this rule, missing values of a stack's daily PM mass rate are imputed using the monthly mean PM mass rate of the stack's treatment group. We add fixed effects for the different types of abatement devices which an industry has across all of its stacks (Type 1, Type 2, or both). Type 1 Devices are the omitted level of device type fixed effect. Approximately 80%, 10%, and 10% of plants are set to Type 1, Type 2, and having both types, respectively. All columns control for plant characteristics including capital expenditure, operating cost, log(total heat output), mean boiler installation year, and their corresponding indicators for missing values. In addition to plant controls, columns 2, 4, 6, and 8 add year-month fixed effects to control for time variant differences common in each plant. We also apply the inverse probability weighting method in columns 3 and 4. The probability of reporting in a month is predicted using a probit model where the only explanatory variable is an indicator variable for the treatment status in a prior experiment that randomized CEMS installation timing. Robust standard errors in parentheses are clustered at the plant level with statistical significance indicated by $^{*}p<0.10$; $^{**}p<0.05$; $^{***}p<0.01$.

F.8. Prior Regulations in Other Markets

The market experiment in this paper is layered on top of existing regulations which mandate investment in abatement and monitoring technologies such as CEMS devices. This layered mandate is typical of the way in which markets have been implemented throughout the world and studied throughout the economics literature. In Table F9 we canvas 6 different emissions markets in the US and EU, give the prior regulations which plants were facing, whether those regulations were lifted once the market was put in place, and also give a citation to a paper studying those markets.

The table summarizes command-and-control regulations that were in place before or alongside major emissions markets. Each row considers one market. The "Program" column names the emissions market and the pollutant it covers. The "Prior Regulation" column lists relevant regulation which was introduced before or concurrent to the emissions market. CAAA refers to the Clean Air Act Amendments; NSPS refers to New Source Performance Standards (applies to all new sources of NOx and SO2, applying uniform national standard based on best adequately demonstrated technology); BACT refers to Best Available Control Technology (applies to all new sources of NOx and SO2 emitting significant amounts in attainment areas and is at least as strict as NSPS); RACT refers to Best Reasonably Available Control Technology. The "Prior Regulation Lifted" column indicates whether (or how) the prior regulations were adjusted at the point of introducing the emissions market. The "Paper on Market" column gives citations to papers studying the market.

Table F9. Prior regulations across emissions markets

Program (Pollutant)	Location	Year Instituted	Prior Regulation	Prior Regulation Lifted?	Paper on Market
Nitrogen Oxides Budget Program (NOx)	Eastern US	2003	• 1970 CAAA: NSPS, BACT • 1990 CAAA: Required subset of boilers in extreme areas to transition to low-pollution fuel • 1990 CAAA: Acid Rain program manted installing NOx monitoring and abatement normally met with control technology • In 1995 Ozone Transport Commission required existing sources to meet RACT limits and ran OTC NOx budget program	No	Deschênes, Greenstone and Shapiro (2017)
AB-32 cap-and-trade system (CO2)	California	2013	• Bill establishing market included complementary programs and modifications of existing programs, including LCFS, RPS, and efficiency mandates • LCFS adopted in 2009 and "require[d] the carbon intensity of transportation fuels to be reduced by at least ten percent in 2020" (ARB, 48). "In April 2011 California adopted a 33 percent RPS" (Appendix, 21). The efficiency mandates varied by different buildings and appliances (ARB, 37).	No, introduced concurrently	Borenstein et al. (2019)
RECLAIM: Regional Clean Air Incentives Market (NOx)	South Coast Air Basin in Southern California	1994	• 1970 CAAA: NSPS, BACT • In 1990 South Coast Air Basin was only nonattainment area for NOx emissions. Thus, the NOx emissions standards may have been more stringent here than other areas. • Prior to RECLAIM, command-and-control program emphasized advanced control technologies, which they adopted in late 1989.	Relaxed	Fowlie, Holland and Mansur (2012)
Acid Rain Program (SO2)	US	1990	• 1970 CAAA: NSPS, BACT • 1977 CAAA: New coal plants must operate with scrubbers and achieve a certain reduction in potential SO2 emissions (70-90%)	No	Joskow and Schmalensee (1998)
Regional Greenhouse Gas Initiative (CO2)	Northeast US	2009	• All states except Virginia had implemented an RPS • Complementary measures as part of RGGI separate from its emissions cap: RGGI auction revenues used for "energy efficiency purposes". • CAAA in force: favoring substitution to cleaner sources from coal powered generation	No	Murray and Maniloff (2015); Kim and Kim (2016)
European Union Emissions Trading System (CO2)	EU	2005	• Directive 2001/77/EC in 2001 set country specific targets for adoption of renewable energy production. • Directive 2003/30/EC in 2003 promoted biofuels for EU transport.	No	Deschênes, Greenstone and Shapiro (2017)

G. APPENDIX: BENEFIT-COST ANALYSIS

We conduct a benefit-cost analysis of introducing an expanded ETS in Surat covering all plants that burn solid fuel. The analysis compares the social benefits of cleaner air, as measured by the valuation of the additional life-years that would be gained from pollution abatement, against the costs of emissions abatement and monitoring. For this exercise we assume that the ETS is expanded with the cap proportionately scaled to maintain the same regulatory stringency per plant as in the experiment. Table VI summarizes the analysis we describe below.

G.1. Costs of monitoring and abatement

The costs of the ETS include both the monitoring infrastructure necessary for the market and the abatement costs, or cost savings, induced by the market. In the experiment, both treatment and control groups purchased CEMS but these devices were not used under the status quo.

We estimate the annual costs of operating a CEMS system at approximately USD 5000 per plant. We arrive at this number by assuming an annualized capital cost of CEMS of INR 200,000, annual device calibration costs of INR 30,000, annual fees for software licenses and maintenance contracts of INR 60,000, and miscellaneous costs (replacement parts, labor etc) at INR 50,000. The annualized CEMS costs are based on an assumed system cost of INR 800,000 with a 4 year equipment life and no discounting. This equipment life describes the realized experience of some plants in our sample but is lower than typical manufacturer claims. License fee and contract costs are based on conversations with vendors and industry. Calibration costs assume three visits a year.

Partly offsetting this monitoring cost, our estimates imply a reduction in abatement costs of roughly USD 650 per plant-year, despite that treatment plants are operating at a sharply lower level of emissions than control plants (row A2). The net per plant costs of monitoring are therefore reduced to closer to USD 4,000. There were a total of 906 registered solid fuel burning plants in Surat during the period of the market and thus in a hypothetical scale-up to cover all plants, we estimate the total private costs, inclusive of both monitoring and abatement, to be USD 3.91 million per year.

G.2. Benefits of lower pollution

The benefit of the ETS is cleaner air. We monetize the benefit of cleaner air by using estimates of the damage from particulates, in terms of life-years gained, and valuing these life-years using estimates of the value of statistical life.

The first step is to estimate how much ETS would reduce ambient pollution (as opposed to industrial pollution emissions). This step is non-trivial because there are many sources of PM2.5. A simple estimate of the impact of the ETS is that ambient pollution would fall by an amount equal to the percentage reduction in emissions due to the regulation, multiplied by the total contribution of these sources to ambient concentrations.

The first term is simply the assumed reduction in emissions, either 10%, 30% or 50%, across columns 1 to 3. For the second term we turn to an estimate from the atmospheric science literature that industrial sources in Surat raise ambient fine particle concentrations by 28.32 $\mu g/m^3$. Guttikunda, Nishadh and Jawahar (2019) use pollution inventories combined with an atmospheric dispersion model to apportion ambient particulate concentrations in Indian cities to different sources.[36] The authors estimate annual average ambient PM 2.5 concentrations in the city at 88.5 $\mu g/m^3$, with 32% (or 28.32 $\mu g/m^3$) coming from local industry. Then the Surat ETS applied to all plants in the city would reduce fine particulate pollution by $0.30 \times 28.32 = 8.5 \mu g/m^3$ (panel B, column 2).

The second step is to estimate the life-years gained from lower pollution. A large literature has attempted to quantify the impact of air pollution on life expectancy. Ebenstein et al. (2017) use a spatial regression discontinuity, at high levels of pollution in China, to estimate that a 10 $\mu g/m^3$ reduction in pollution results in a 0.98 year increase in life expectancy. Other estimates in the literature include 0.61 years (Pope, Ezzati and Dockery, 2009) and 0.12 years (calculated from Table S2 in Apte et al. (2018)).

These estimates should be interpreted as the benefits of long-run changes in pollution. If we were to assume that an ETS were implemented in Surat for 70 years (roughly the current life

[36] Their updated assessment for Surat is available at: `https://urbanemissions.info/india-apna/surat-india/`.

expectancy in India), reducing pollution each year by 8.5 $\mu g/m^3$, then the health benefits from Ebenstein et al. (2017) would suggest life expectancy gains of $0.98 \times 8.5/10 = 0.83$ years per person. The population of Surat in 2021 as estimated 7.5 million people. Thus the total gain in life years would be these per-person estimates multiplied by the city population, or 6.24 million years. Assuming these accrue gradually over the 70 year period of the ETS, the gain from a single year of the program would be 89,208 years.

The third step is to value the life-years gained. We use a VSL estimate for India of USD 665,000 (Nair et al., 2021) and apply this equally to every year of an assumed 70 year life yielding a dollar value of USD 9,500 per life-year gained. This number, combined with the life years gained from a year of the ETS, would imply a single year health benefit of USD 847 million and thus a benefit to cost ratio as high as 215 to 1 (panel E, row 1, column 2). Using the lower estimates of health benefits from Apte et al. (2018) yields a benefit to cost ratio of 26 (panel E, row 4, column 2). By either estimate, the benefits of the expanded ETS greatly exceed the total of monitoring and abatement costs.

References cited within Appendices

Apte, Joshua S., Michael Brauer, Aaron J. Cohen, Majid Ezzati, and C. Arden Pope III. 2018. "Ambient PM2.5 Reduces Global and Regional Life Expectancy." *Environmental Science and Technology Letters*, 5(9): 546–551.

Borenstein, Severin, James Bushnell, Frank A. Wolak, and Matthew Zaragoza-Watkins. 2019. "Expecting the Unexpected: Emissions Uncertainty and Environmental Market Design." *American Economic Review*, 109(11): 3953–3977.

Deschênes, Olivier, Michael Greenstone, and Joseph S. Shapiro. 2017. "Defensive Investments and the Demand for Air Quality: Evidence from the NOx Budget Program." *American Economic Review*, 107(10): 2958–2989.

Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan. 2018. "The Value of Regulatory Discretion: Estimates From Environmental Inspections in India." *Econometrica*, 86(6): 2123–2160.

Ebenstein, Avraham, Maoyong Fan, Michael Greenstone, Guojun He, and Maigeng Zhou. 2017. "New Evidence on the Impact of Sustained Exposure to Air Pollution on Life Expectancy from China's Huai River Policy." *Proceedings of the National Academy of Sciences of the United States of America*, 114(39): 10384–10389.

Fowlie, Meredith, Stephen P. Holland, and Erin T. Mansur. 2012. "What Do Emissions Markets Deliver and to Whom? Evidence from Southern California's NOx Trading Program." *American Economic Review*, 102(2): 965–993.

Guttikunda, Sarath K., K. A. Nishadh, and Puja Jawahar. 2019. "Air Pollution Knowledge Assessments (APnA) for 20 Indian Cities." *Urban Climate*, 27: 124–141.

Joskow, Paul L., and Richard Schmalensee. 1998. "The Political Economy of Market-Based Environmental Policy: The U.S. Acid Rain Program." *Journal of Law and Economics*, 41: 37–83.

Kim, Man-Keun, and Taehoo Kim. 2016. "Estimating Impact of Regional Greenhouse Gas Initiative on Coal to Gas Switching Using Synthetic Control Methods." *Energy Economics*, 59: 328–335.

Murray, Brian C., and Peter T. Maniloff. 2015. "Why Have Greenhouse Emissions in RGGI States Declined? An Econometric Attribution to Economic, Energy Market, and Policy Factors." *Energy Economics*, 51: 581–589.

Nair, Moorthy, Hemant Bherwani, Shahid Mirza, Saima Anjum, and Rakesh Kumar. 2021. "Valuing Burden of Premature Mortality Attributable to Air Pollution in Major Million-plus Non-attainment Cities of India." *Scientific Reports*, 11: 1–15.

Pope, C. Arden III, Majid Ezzati, and Douglas W. Dockery. 2009. "Fine-Particulate Air Pollution and Life Expectancy in the United States." *New England Journal of Medicine*, 360(4): 376–386.

Sudarshan, Anant. 2023. "Monitoring Industrial Pollution in India." In *Introduction to Development Engineering*. , ed. Temina Madon, Ashok J. Gadgil, Richard Anderson, Lorenzo Casaburi, Kenneth Lee and Arman Rezaee, 161–182. Springer International Publishing.